SPECIAL COLLECTOR'S EDITION

SCIENTIFIC AMERICAN

TOP SCIENCE STORIES OF 2019

Lifesaving workouts

The first picture of a black hole

Dengue vaccine danger?

Future of the moon

A new face on the family tree

Vaping myths exposed

WINTER 2019/2020

© 2019 Scientific American

Scientific American Unlimited

Perfect for science fans, gain access to all of our publications, apps & the full website experience.

Digital archive access back to 1845 including articles by Einstein, Curie and other timeless luminaries in more than 7,000 issues!

12 print and digital issues of *Scientific American* per year

More than 150 eBooks and Collector's Editions

Access to *Scientific American Mind*, *Space & Physics* and *Health & Medicine*

More than 200 articles per month, including news and commentary, on ScientificAmerican.com

sciam.com/unlimited

Scientific American is a registered trademark of Springer Nature America, Inc

SCIENTIFIC AMERICAN®
ESTABLISHED 1845

Top Science Stories of 2019 is published by the staff of Scientific American, with project management by:

Acting Editor in Chief: **Curtis Brainard**
Senior Editor, Collections: **Andrea Gawrylewski**

Creative Director: **Michael Mrak**
Issue Designer: **Lawrence R. Gendron**
Photography Editor: **Monica Bradley**
Assistant Photo Editor: **Liz Tormes**

Copy Director: **Maria-Christina Keller**
Senior Copy Editors: **Angelique Rondeau, Daniel C. Schlenoff, Aaron Shattuck**
Managing Production Editor: **Richard Hunt**
Prepress and Quality Manager: **Silvia De Santis**

Editorial Administrator: **Ericka Skirpan**
Executive Assistant Supervisor: **Maya Harty**

President: **Dean Sanderson**
Executive Vice President: **Michael Florek**
Vice President, Commercial: **Andrew Douglas**
Publisher and Vice President: **Jeremy A. Abbate**
Associate Vice President, Business Development: **Diane McGarvey**
Head, Marketing and Product Management: **Richard Zinken**
Marketing Director, Institutional Partnerships and Customer Development: **Jessica Cole**
Programmatic Product Manager: **Zoya Lysak**
Marketing Manager: **Chris Monello**
Senior Commercial Operations Coordinator: **Christine Kaelin**
Head, Communications, USA: **Rachel Scheer**
Press Manager: **Sarah Hausman**
Director, Integrated Media: **Jay Berfas**
Director, Integrated Media: **Matt Bondlow**
Manager, Global Media Alliances: **Brendan Grier**
Custom Publishing Editor: **Lisa Pallatroni**
Rights and Permissions Manager: **Felicia Ruocco**
Senior Administrator, Executive Services: **May Jung**
Production Controller: **Madelyn Keyes-Milch**
Advertising Production Controller: **Dan Chen**

COVER IMAGES (*clockwise from top left*): BRYAN CHRISTIE (*skeleton*); EVENT HORIZON TELESCOPE [CC BY 4.0 (HTTPS://CREATIVECOMMONS.ORG/LICENSES/BY/4.0)] (*black hole*); NASA (*Buzz Aldrin*); GETTY IMAGES (*vape pen*); SKULL: DALE OMORI AND LIZ RUSSELL (*photograph*); JENNIFER TAYLOR (*composite image*) AND CLEVELAND MUSEUM OF NATURAL HISTORY; GETTY IMAGES (*mosquito*)

FROM THE EDITOR

This Year in Science

Of all the scientific accomplishments in 2019, one in particular stands out: an image of a flaming ring of starlight surrounding a dark circle. Black holes had long been theorized to exist but never directly observed, and scientists had until this year only guessed at what they might look like. But finally, it happened. After more than 10 years of data collection from 11 radio telescopes around the world, astronomers captured the shadow of the gargantuan black hole at the heart of the galaxy Messier 87 (*above*).

Few things inspire as much awe as seeing what was once unseeable, but there were a number of other extraterrestrial milestones in the past 12 months. On January 1, NASA's New Horizons probe flew by the farthest body ever visited by a spacecraft—the snowman-shaped object called Ultima Thule (*page 32*). Local temperature was reportedly 400 degrees Fahrenheit below zero. Then, in July, we celebrated the 50-year anniversary of the *Apollo 11* moon landing. Precious lunar territory may soon be hotly contested by spacefaring nations and commercial interests in search of new resources, Adam Mann writes on page 38.

Back on Earth, incredible achievements abound. Neuroscientists restored circulation and molecular and cellular functions in excised pig brains four hours after death, making them appear relatively normal but for the absence of brain waves indicative of consciousness (*page 28*). Elsewhere in the resurrection department, researchers re-created the scent genes and inhaled the fragrance of plants that have been extinct for decades (*page 68*). As for the origins of life, analyses of ancient coral-like fossils suggest that complex animals got their start millions of years earlier than previously thought (*page 52*).

Every year it seems that science becomes more central to tackling society's problems. With engineers trying to head off dangerous global warming, Richard Conniff takes stock of the technologies available to remove carbon dioxide directly from the atmosphere (*page 84*). A team of investigators working in Guinea-Bissau devised and tested a controversial but enticing hypothesis that some vaccines for specific diseases can confer protection for many other maladies, such as tuberculosis, malaria or flu (*page 20*). Cailin O'Connor and James Owen Weatherall apply network science to explain why misinformation goes viral so easily (*page 110*).

The elegance of science is its ability to be agile and adaptive—it asks how we know what we know, posing questions that are continually refined and reframed. And sometimes we're lucky enough to get signposts that show—through a shadowy ring of fire perhaps—we're on the right path.

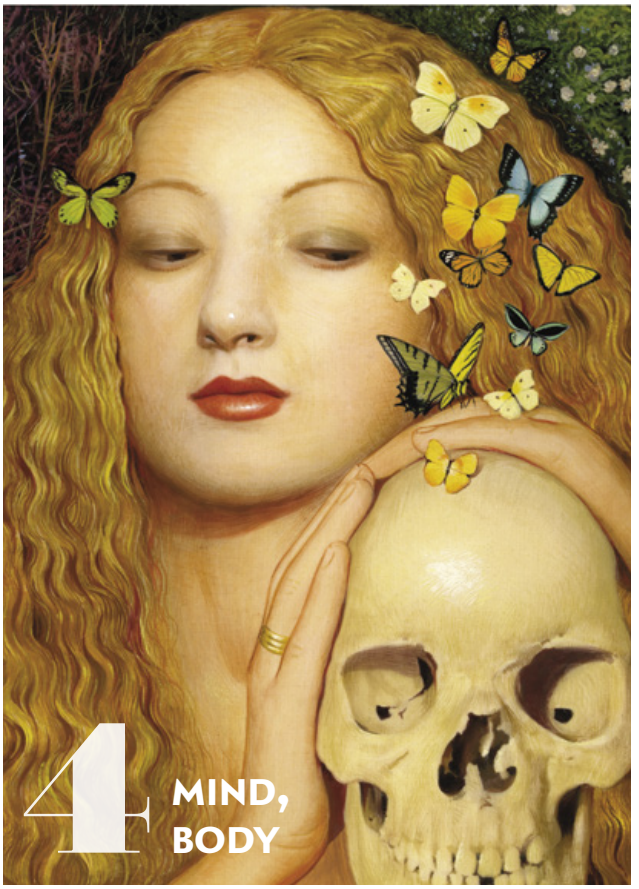
Andrea Gawrylewski
Senior Editor, Collections
editors@sciam.com

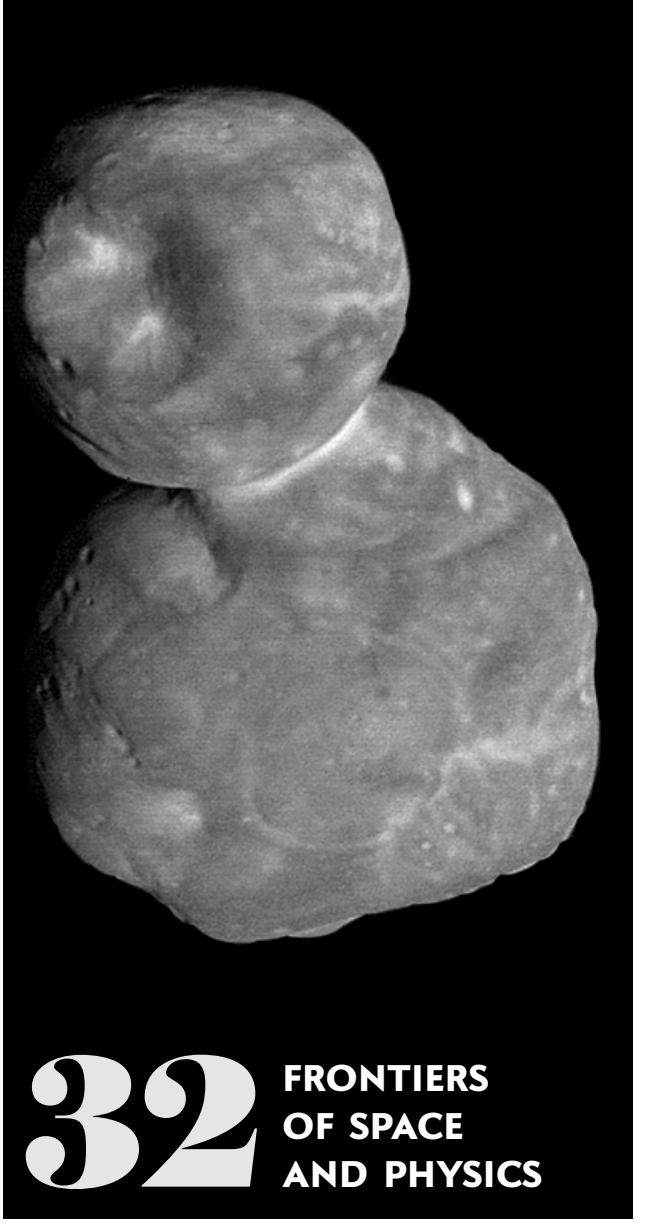
© 2019 Scientific American

SPECIAL EDITION

SCIENTIFIC AMERICAN

Volume 28, Number 5, Winter 2019/2020





4 Face Values
Brain regions that process faces reveal deep insights into the neural mechanisms of vision.
By Doris Tsao

12 The Point of a Period
Age-old taboos against menstruation have led to a lack of research on how women's cycles work, with serious consequences for their health.
By Virginia Sole-Smith

20 Vaccines Reimagined
A controversial theory holds that one immunization, given properly, can protect against many diseases besides its target.
By Melinda Wenner Moyer

28 Is Death Reversible?
An experiment that partially revived slaughterhouse pig brains raises questions about the precise end point of life.
By Christof Koch

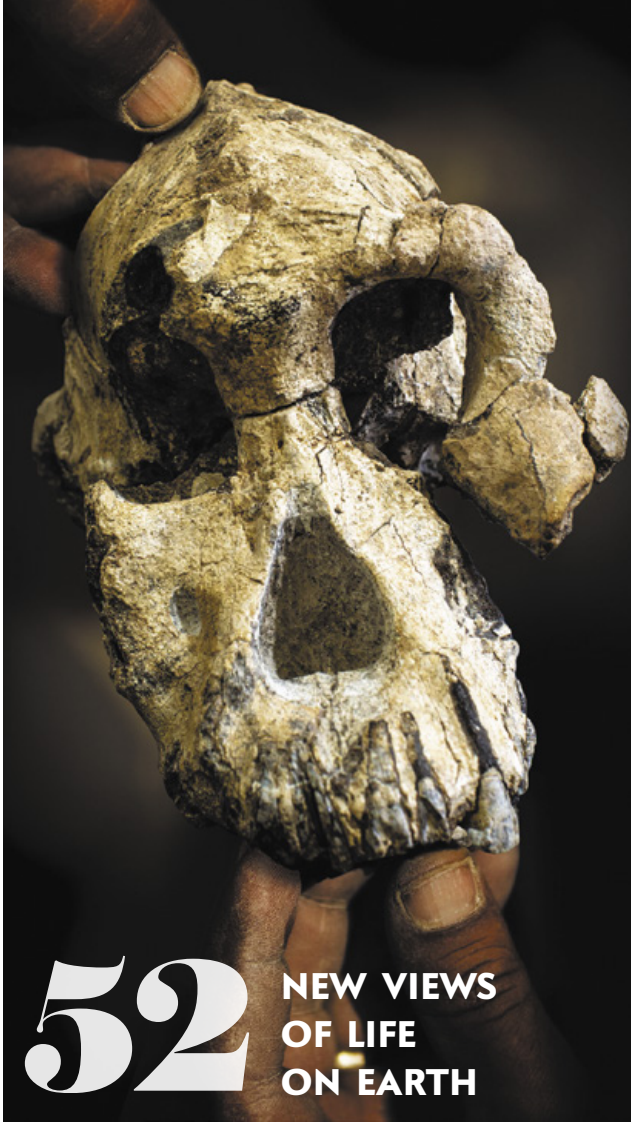
32 Distant New World
Ultima Thule, the farthest object ever visited by a spacecraft, is revealing our solar system's deepest history—and, just maybe, revolutionizing planetary science.
By Corey S. Powell

36 Finally, the Face of a Black Hole
After more than a decade of effort, a global network of radio telescopes has revealed the first ever picture of an enigmatic hole in spacetime.
By Seth Fletcher

38 Lunar Land Grab
A new race could be heating up to claim valuable moon terrain amid uncertain laws.
By Adam Mann

44 Crystals in Time
Surprising new states of matter called time crystals show the same symmetry properties in time that ordinary crystals do in space.
By Frank Wilczek

© 2019 Scientific American



52 NEW VIEWS OF LIFE ON EARTH

82 SCIENCE MEETS SOCIETY

52 The Rise of Animals
New fossils and analyses of ancient ocean chemistry reveal the surprisingly deep roots of the Cambrian explosion.
By Rachel A. Wood

60 Evolved to Exercise
Unlike our ape cousins, humans require high levels of physical activity to be healthy.
By Herman Pontzer

68 Ghost Flowers
The genes of Hawaiian plants, extinct for more than a century, have been brought back from the dead. Today we can smell their scents.
By Rowan Jacobsen

78 A Face from Deep Time
A long-sought fossil cranium could redraw the human family tree.
By Kate Wong

82 Air Inequality
U.S. racial minorities are exposed to more air pollution than white people yet cause less of it.
By Andrea Thompson

84 The Last Resort
Can we remove enough CO_2 from the atmosphere to slow or even reverse climate change? *By Richard Conniff*

92 The Dengue Debacle
Is a runaway immune reaction making a dengue vaccine dangerous? *By Seema Yasmin and Madhusree Mukerjee*

102 Divide or Conquer
Five nations are asserting rights to vast, overlapping portions of the Arctic Ocean seafloor.
By Mark Fischetti

106 Setting Boundaries
By Katie Peek and Mark Fischetti

110 Why We Trust Lies
The most effective misinformation starts with seeds of truth.
By Cailin O'Connor and James Owen Weatherall

DEPARTMENTS

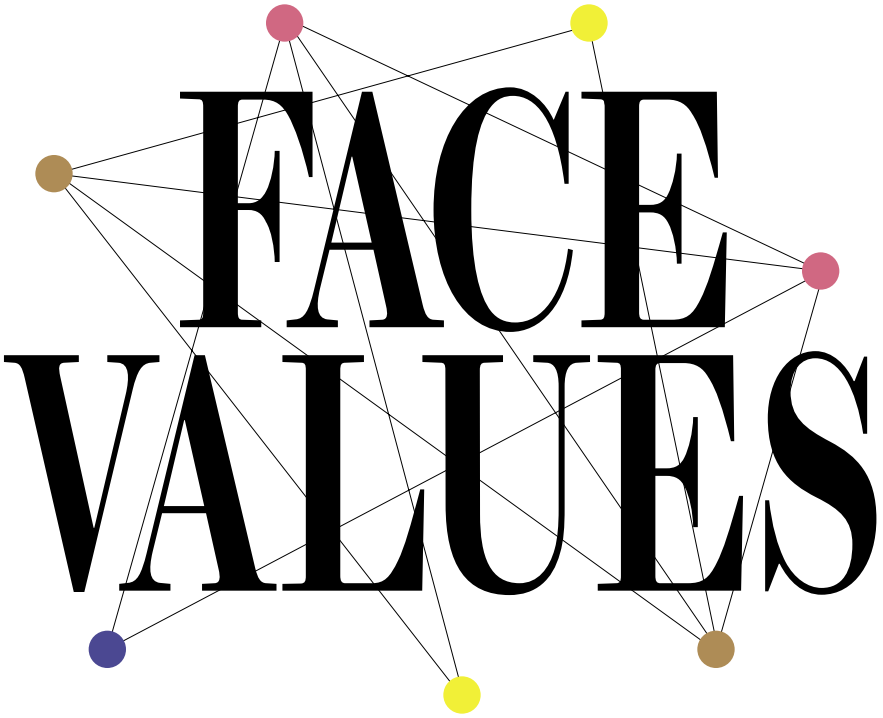
1 FROM THE EDITOR
This Year in Science

116 END NOTE
Smoke Screen
Social media bots promote unproved benefits of e-cigarettes. *By Mark Fischetti and Sree Priyanka Uppu*

Articles in this special issue are updated or adapted from previous issues of *Scientific American* and *Nature* and from ScientificAmerican.com.
Copyright © 2019 Scientific American, a division of Springer Nature America, Inc. All rights reserved. Scientific American Special (ISSN 1936-1513), Volume 28, Number 5, Winter 2019/2020, published by Scientific American, a division of Springer Nature America, Inc., 1 New York Plaza, Suite 4600, New York, N.Y. 10004-1562. Canadian BN No. 127387652RT; TVQ1218059275 TQ0001. To purchase additional quantities: U.S., $13.95 each; elsewhere, $17.95 each. Send payment to Scientific American Back Issues, P.O. Box 3187, Harlan IA 51537. Inquiries: fax 212-355-0408 or telephone 212-451-8415. Printed in U.S.A.

© 2019 Scientific American

© 2019 Scientific American



FACE VALUES

Brain regions that process faces reveal deep insights into the neural mechanisms of vision

By Doris Y. Tsao

WHEN I WAS IN HIGH SCHOOL, I LEARNED ONE DAY ABOUT THE density of curves in an introductory course on calculus. A simple pair of differential equations, which model the interactions of predators and prey, can give rise to an infinite number of closed curves—picture concentric circles, one nested within another, like a bull's-eye. What is more, the density of these curves varies depending on their location.

This last fact seemed so strange to me. I could easily imagine a finite set of curves coming close together or pulling apart. But how could an infinity of curves be denser in one region and less dense in another? I soon learned that there are different types of infinity, which have paradoxical qualities, like Hilbert's Hotel (where the rooms are always fully booked but new guests can always be accommodated) and the Banach-Tarski apple (which can be split into five pieces and rearranged to make two apples of equal volume as the original). I spent hours poring over these mathematical proofs. Ultimately they struck me as symbolic magic of no real consequence, but the seed of interest had taken root.

Illustration by Brian Stauffer

© 2019 Scientific American

Later, as an undergraduate at the California Institute of Technology, I learned about the experiments of David Hubel and Torsten Wiesel and their landmark discovery of how a region in the brain called the primary visual cortex extracts edges from the images relayed from the eyes. I realized that what had mystified me back in high school was the act of trying to *imagine* different densities of infinity. Unlike the mathematical tricks I had studied in high school, the edges that Hubel and Wiesel described are processed by neurons, so they actually exist in the brain. I came to recognize that visual neuroscience was a way to understand how this neural activity gives rise to the conscious perception of a curve.

The sense of excitement this realization triggered is hard to describe. I believe at each stage in life one has a duty. And the duty of a college student is to dream, to find the thing that captures one's heart and seems worth devoting a whole life to. Indeed, this is the single most important step in science—to find the right problem. I was captivated by the challenge of understanding vision and embarked on a quest to learn how patterns of electrical activity in the brain are able to encode perceptions of visual objects—not just lines and curves but even objects as hard to define as faces. Accomplishing this objective required pinpointing the specific brain regions dedicated to facial recognition and deciphering their underlying neural code—the means by which a pattern of electrical impulses allows us to identify people around us.

The journey of discovery began in graduate school at Harvard University, where I studied stereopsis, the mechanism by which depth perception arises from differences between the images in the two eyes. One day I came across a paper by neuroscientist Nancy Kanwisher, now at the Massachusetts Institute of Technology, and her colleagues, reporting the discovery of an area in the human brain that responded much more strongly to pictures of faces than to images of any other object when a person was inside a functional magnetic resonance imaging (fMRI) brain scanner. The paper seemed bizarre. I was used to the brain being made of parts with names like basal ganglia and orbitofrontal cortex that had some vague purpose one could only begin to fathom. The concept of an area specifically devoted to processing faces seemed all too comprehensible and therefore impossible. Anyone could make a reasonable conjecture about the function of a face area—it should probably represent all the different faces that we know and something about their expression and gender.

As a graduate student, I had used fMRI on monkeys to identify areas activated by the perception of three-dimensionality in images. I decided to show pictures of faces and other objects to a monkey. When I compared activation in the monkey's brain to faces with activation to other objects, I found several areas that lit up selectively to faces in the temporal lobe (the area underneath the temple)—specifically in a region called the inferotemporal (IT) cortex. Charles Gross, a pioneer in the field of object vision, had discovered face-selective neurons in the IT cortex of macaques in the early 1970s. But he had reported that these cells were randomly scattered throughout the IT cortex. Our fMRI results provided the first indication that face cells might be concentrated into defined regions.

FACE PATCHES

AFTER PUBLISHING MY WORK, I was invited to give a talk describing the fMRI study for a faculty position at Caltech, but I was not offered the job. Many people were skeptical of the value of fMRI, which measures local blood flow, the brain's plumbing. They argued that showing increased blood flow to a brain area when a subject is looking at faces falls far short of clarifying what neurons in the area are actually encoding because the relation between blood flow and electrical activity is unclear. Perhaps by chance these face patches simply contained a slightly larger number of neurons responsive to faces, like icebergs randomly clustered at sea.

Because I had done the imaging experiment in a monkey, I could directly address this concern by inserting an electrode into an fMRI-identified face area and asking, What images drive single neurons in this region most strongly? I performed this experiment together with Winrich Freiwald, then a postdoctoral fellow in Margaret Livingstone's laboratory at Harvard, where I was a graduate student. We presented faces and other objects to a monkey while amplifying the electrical activity of individual neurons recorded by the electrode. To monitor responses in real time, the neurons' electrical signals were converted to an audio signal that we could hear with a loudspeaker in the lab.

This experiment revealed an astonishing result: almost every single cell in the area identified through fMRI was dedicated to processing faces. I can recall the excitement of our first recording, hearing the "pop" of cell after cell responding strongly to faces and very little to other objects. We sensed we were on to something important, a piece of cortex that could reveal the brain's high-level code for visual objects. Marge remarked on the face patches: "You've found a golden egg."

I also remember feeling surprised during that first experiment. I had expected the face area would contain cells that responded selectively to specific individuals, analogous to orientation-selective cells in the primary visual cortex that each respond to a specific edge orientation. In fact, a number of well-publicized studies had suggested that single neurons can be remarkably selective for the faces of familiar people—responding, say, only to Jennifer Aniston. Contrary to my expectation, each cell seemed to fire vigorously to almost any face.

I plugged madly away at Photoshop during these early experiments and found that the cells responded not just to faces of humans and monkeys but even to highly simplified cartoon faces.

IN BRIEF

Understanding vision remains one of the grand challenges that neuroscientists confront.

One key aspect of this problem relates to the way the brain identifies faces, the most important social emblem.

Neurons in defined sections of the cerebral cortex, called face patches, are dedicated to recognizing faces.

Uncovering the organization of the face-patch system served as a prelude to deducing the underlying computations that the brain makes to identify faces.

This neural code may serve as a Rosetta stone for representing other objects besides faces.

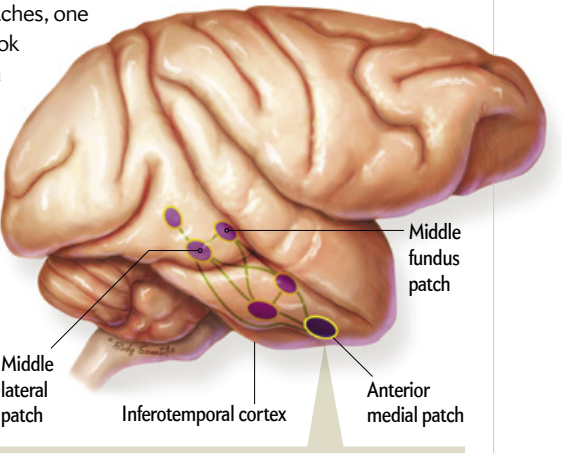
© 2019 Scientific American

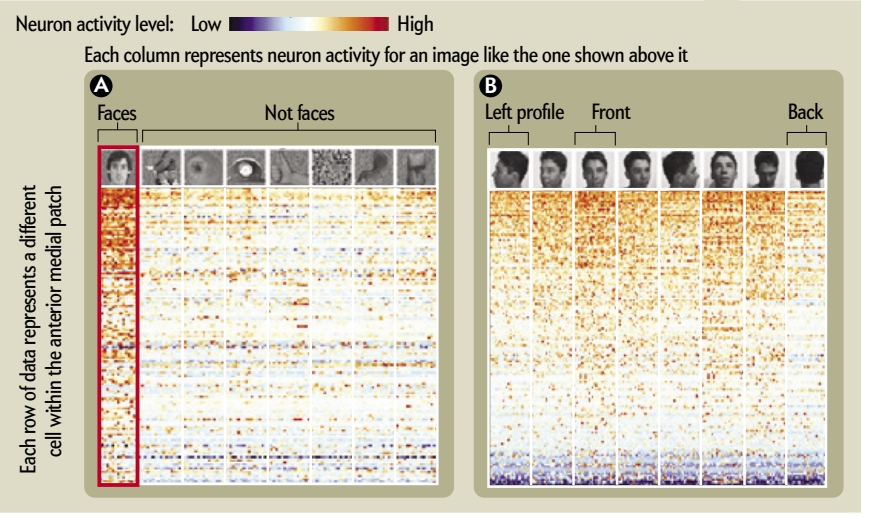
Observing this phenomenon, I decided to create cartoon faces with 19 different features that seemed pertinent to defining the identity of a face, including inter-eye distance, face aspect ratio and mouth height, among other characteristics. We then went on to alter these values—the inter-eye distance, for instance, varied from being almost cyclopean to just inside the face boundary. Individual cells responded to most faces but interestingly did not always exhibit the exact same rate of firing to all faces. Instead there was a systematic variation in their response: when we plotted the firing of cells for the different cartoon features, we found a pattern in which there was a minimal response to one feature extreme—the smallest inter-eye distance, for instance—and a maximal response to the opposite extreme—the largest eye separation—with intermediate responses to feature values in the middle. The response as a function of the value for each feature looked like a ramp, a line slanted up or down.

Once again, I was invited to give a job talk at Caltech. Returning, I had more to offer than just fMRI images. With the addition of the new results from single-cell recordings, it was clear to everyone that these face patches were real and likely played an important role in facial recognition. Furthermore, understanding their underlying neural processes seemed like an effective way to gain traction on the general problem of how the brain represents visual objects. This time I was offered the job.

Where Are the Face Detectors?

A set of six nodes in the inferotemporal cortex of both brain hemispheres specializes in identifying faces. These "face patches" function as an assembly line: in the middle lateral and middle fundus patches, one neuron might become active when faces look straight ahead; another might turn on for a face looking to the right. At the end of the assembly line, in the anterior medial patch, varying views are stitched together. Neurons in this patch are active in response to the face of a specific individual, no matter if the view is from the front or side. Responses from a face patch of one monkey are generated for faces but not objects (*red areas in* A) and for the same individual, such as the dark-haired man, from varying angles (*red areas in* B).




Neuron activity level: Low High
Each column represents neuron activity for an image like the one shown above it
A
Faces
Not faces
B
Left profile
Front
Back
Each row of data represents a different cell within the anterior medial patch

CONTRAST IS KEY

AT CALTECH, my colleagues and I dug deeper into the question of how these cells detect faces. We took inspiration from a paper by Pawan Sinha, a vision and computational neuroscientist at M.I.T., that suggested faces could be discerned by checking for specific contrast relations between different regions of the face—whether the forehead region is brighter than the mouth region, for example. Sinha suggested a clever way to determine *which* contrast relations can be used to recognize a face: they should be the ones that are immune to changes in lighting. For example, left-eye-darker-than-nose is a useful feature for detecting a face because it does not matter if a face is photographed with lighting from above, left, right or below: the left eye is *always* darker than the nose (check for yourself).

From a theoretical standpoint, this idea provides a simple, elegant computational mechanism for facial recognition, and we wondered whether face cells might be using it. When we measured the response of cells to faces in which different regions varied in brightness, we found that cells often had a significant preference for a particular contrast feature in an image.

To our astonishment, almost all the cells were wholly consistent in their contrast preferences—just a single cell was found that preferred the opposite polarity. Moreover, the preferred features were precisely those identified by Sinha as being invulnerable to lighting changes. The experiment thus confirmed that face cells use contrast relations to detect faces.

More broadly, the result confirmed that these cells truly were face cells. At talks, skeptics would ask, How do you know? You can't test every possible stimulus. How can you be sure it's a face cell and

FROM "FUNCTIONAL COMPARTMENTALIZATION AND VIEWPOINT GENERALIZATION WITHIN THE MACAQUE FACE-PROCESSING SYSTEM," BY WINRICH A. FREIWALD AND DORIS Y. TSAO, IN *SCIENCE*, VOL. 330; NOVEMBER 5, 2010 (*data grids* and *photo insets*)

Illustration by Body Scientific (brain)

© 2019 Scientific American

not a pomegranate cell or a lawn mower cell? This result nailed it for me. The precise match between the way cells reacted to changes in contrast between different parts of the face and Sinha's computational prediction was uncanny.

Our initial experiments had revealed two nearby cortical patches that lit up to faces. But after further scanning (with the help of a contrast agent that increased severalfold the robustness of the signal), it became clear that there are in fact six face patches in each of the brain's two hemispheres (making a dozen golden eggs total). They are distributed along the entire length of the temporal lobe. These six patches, moreover, are not randomly scattered throughout the IT cortex. They are located in similar locations across hemispheres in each animal. Moreover, work by our group and others has found that a similar pattern of multiple face patches spanning the IT cortex exists in humans and other primates such as marmosets.

This observation of a stereotyped pattern suggested that the patches might constitute a kind of assembly line for processing faces. If so, one would expect the six patches to be connected to one another and each patch to serve a distinct function.

To explore the neural connections among patches, we electrically stimulated different patches with tiny amounts of current—a technique called microstimulation—while the monkey was inside an fMRI scanner. The goal was to find out what other parts of the brain light up when a particular face patch is stimulated. We discovered that whenever we stimulated one face patch, the other patches would light up but not the surrounding cortex, indicating that, indeed, the face patches are strongly interconnected. Furthermore, we found that each patch performs a different function. We presented pictures of 25 people, each at eight different head orientations, to monkeys and recorded responses from cells in three regions: the middle lateral and middle fundus patches (ML/MF), the anterior lateral patch (AL) and the anterior medial patch (AM).

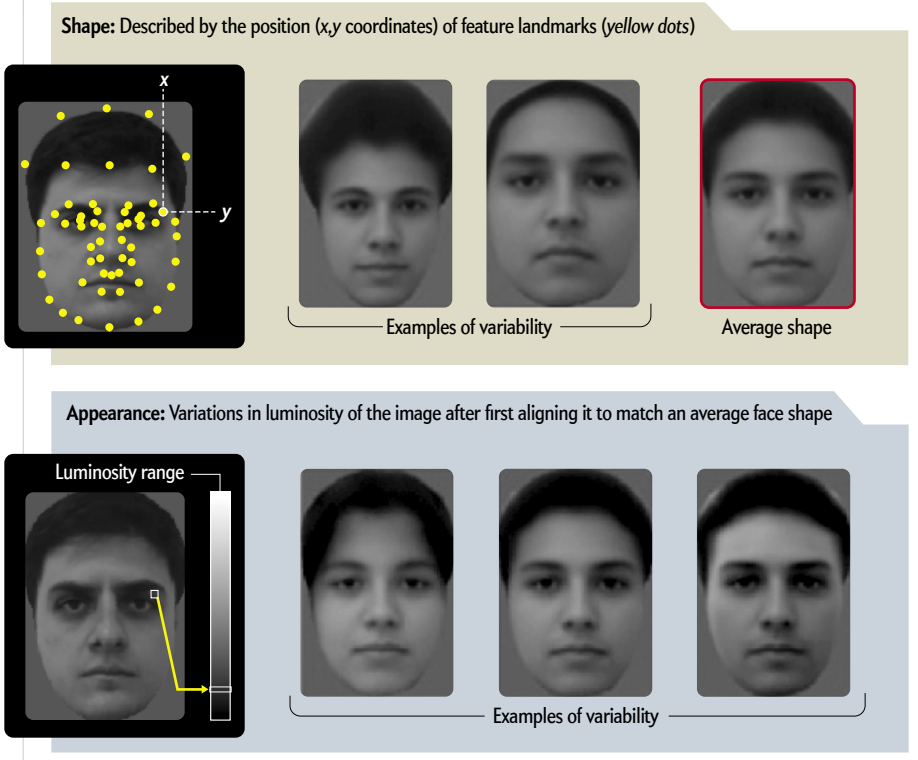
We found striking differences among these three regions. In ML/MF, cells responded selectively to specific views. For example, one cell might prefer faces looking straight ahead, whereas another might opt for faces looking to the left. In AL, cells were less view-specific. One class of cells responded to faces looking up, down and straight ahead; another responded to faces looking to the left or right. In AM, cells responded to specific individuals regardless of whether the view of the face was frontal or in profile. Thus, at the end of the network in AM, view-specific representations were successfully stitched into a view-invariant one.

Apparently face patches do act as an assembly line to solve one of the big challenges of vision: how to recognize things around us despite changes in the way they look. A car can have any make and color, appear at any viewing angle and distance, and be partially obscured by closer objects such as trees or other cars. Recognizing an object despite these visual transformations is called the invariance problem, and it became clear to us that a major function of the face-patch network is to overcome this impediment.

Given the great sensitivity of cells in face patches to changes in facial identity, one might expect that altering these cells' responses

Shape + Appearance = Face

Identifying the face patches was only a first step. It then became necessary to explore what happens in the neurons within each patch, setting off a search for the brain's coding scheme for faces. To derive quantitative measures for faces, the Tsao laboratory came up with 25 features for shape and 25 for appearance that could be used by each neuron in a face patch—a 50-dimensional face space. The shape features can be thought of as those defining the skeleton—how wide the head is or the distance between the eyes. The appearance features specify the face's surface texture (complexion, eye or hair color, and so on).


Shape: Described by the position (x,y coordinates) of feature landmarks (yellow dots)
x
y
Examples of variability
Average shape
Appearance: Variations in luminosity of the image after first aligning it to match an average face shape
Luminosity range
Examples of variability

FROM "THE CODE FOR FACIAL IDENTITY IN THE PRIMATE BRAIN," BY LE CHANG AND DORIS Y. TSAO, IN *CELL*, VOL. 169, NO. 6; JUNE 1, 2017 (*face images*)

Graphics by Jen Christiansen

© 2019 Scientific American

should modify an animal's perception of facial identity. Neuroscientists Josef Parvizi and Kalanit Grill-Spector of Stanford University had electrically stimulated a face-patch area in human subjects who had electrodes implanted in their brains for the purpose of identifying the source of epileptic seizures and found that stimulation distorted the subjects' perception of a face.

We wondered whether we would find the same effect in monkeys when we stimulated their face patches. Would doing so alter the perception only of faces, or would it affect that of other objects as well? The boundary between a face and a nonface object is fluid—one can see a face in a cloud or an electrical outlet if prompted. We wanted to use electrical microstimulation as a tool to delineate precisely what constitutes a face for a face patch. We trained monkeys to report whether two sequentially presented faces were the same or different. Consistent with the earlier results in humans, we found that microstimulation of face patches strongly distorted perception so that the animal would always signal two identical faces as being different.

Interestingly, microstimulation had no effect on the perception of many nonface objects, but it did significantly affect responses to a few objects whose shape is consistent with a face—apples, for one. But why does this stimulation influence the perception of an apple?

One possibility is that the face patches are typically used to represent not just faces but also other round objects like apples. Another hypothesis is that face patches are not normally used to represent these objects, but stimulation induces an apple to appear facelike. It remains unclear whether face patches are useful for detecting any nonface objects.

CRACKING THE CODE

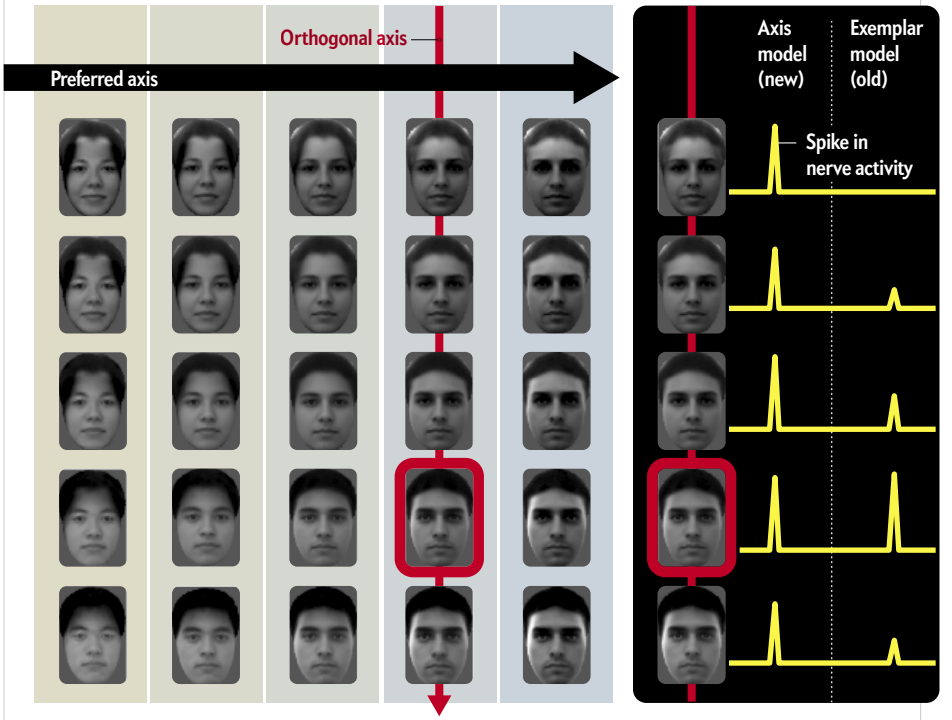
UNCOVERING the organization of the face-patch system and properties of the cells within was a major accomplishment. But my dream when we first began recording from face patches was to achieve something more. I had intuited that these cells would allow us to crack the neural code for facial identity. That means understanding how individual neurons process faces at a level of detail that would let us predict a cell's response to any given face or decode the identity of an arbitrary face based only on neural activity.

The central challenge was to figure out a way to describe faces quantitatively with high precision. Le Chang, then a postdoc in my lab, had the brilliant insight to adopt a technique from the field of computer vision called the active appearance model. In this approach, a face has two sets of descriptors, one for shape and another for appearance. Think of the shape features as those defined by the skeleton—how wide the head is or the distance between the eyes. The appearance features define the surface texture of the face (complexion, eye or hair color, and so on).

To generate these shape and appearance descriptors for faces, we started with a large database of face images. For each face, we placed a set of markers on key features. The spatial locations of these markers described the shape of the face. From these varied shapes, we calculated an average face. We then morphed each face image in the database so its key features exactly matched that of the average face. The resulting im-

The Face Code, at Last

Having 50 coordinates that describe shape and appearance allows for a description of neurons' firing in response to a particular face—a description that functions as a code that can be visualized geometrically. In this code, each face cell receives inputs for a face in the form of the 50 coordinates, or dimensions. The neuron then fires with a particular intensity in response to a certain face (*red outlines*), along what is called the preferred axis. The intensity increases steadily (monotonically) along the preferred axis. Furthermore, the response is the same for every face on an axis at right angles to the preferred axis, even though those faces may look very different. This axis model of facial coding differs from a previous exemplar model that suggests that each neuron fires with maximum intensity to a single most preferred face.

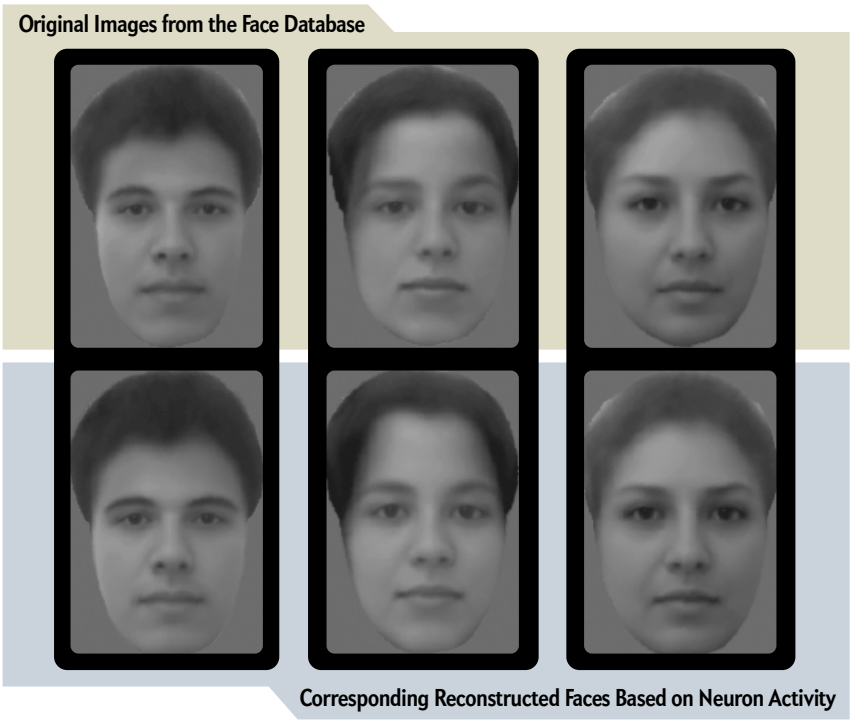


FROM "THE CODE FOR FACIAL IDENTITY IN THE PRIMATE BRAIN," BY LE CHANG AND DORIS Y. TSAO, IN *CELL*, VOL. 169, NO. 6; JUNE 1, 2017 (*face grid*)

© 2019 Scientific American

Pictures Worth 205 Neurons

For a given face, we can predict how a cell will respond by taking a weighted sum of all 50 face coordinates. To predict what face the monkey saw from neuronal activity, this entire process can be reversed: by knowing the response of 205 face cells, it is possible to predict the 50 coordinates defining the exact facial features—and make a highly accurate reconstruction of a given face.



ages constituted the appearance of the faces independent of shape.

We then performed principal components analysis independently on the shape and appearance descriptors across the entire set of faces. This is a mathematical technique that finds the dimensions that vary the most in a complex data set.

By taking the top 25 principal components for shape and the top 25 for appearance, we created a 50-dimensional face space. This space is similar to our familiar 3-D space, but each point represents a face rather than a spatial location, and it comprises much more than just three dimensions. For 3-D space, any point can be described by three coordinates (*x,y,z*). For a 50-D face space, any point can be described by 50 coordinates.

In our experiment, we randomly drew 2,000 faces and presented them to a monkey while recording cells from two face patches. We found that almost every cell showed graded responses—resembling a ramp slanting up or down—to a subset of the 50 features, consistent with my earlier experiments with cartoon faces. But we had a new insight about why this is important. If a face cell has ramp-shaped tuning to different features, its response can be roughly approximated by a simple weighted sum of the facial features, with weights determined by the slopes of the ramp-shaped tuning functions. In other words:

response of face cells = weight matrix × 50 face features

We can then simply invert this equation to convert it to a form that lets us *predict* the face being shown from face cell responses:

50 face features = (1/weight matrix) × response of face cells

At first, this equation seemed impossibly simple to us. To test it, we used responses to all but one of the 2,000 faces to learn the weight matrix and then tried to predict the 50 face features of the excluded face. Astonishingly, the prediction turned out to be almost indistinguishable from the actual face.

A WIN-WIN BET

AT A MEETING in Ascona, Switzerland, I presented our findings on how we could reconstruct faces using neural activity. After my talk, Rodrigo Quian Quiroga, who discovered the famous Jennifer Aniston cell in the human medial temporal lobe in 2005 and is now at the University of Leicester in England, asked me how my cells related to his concept that single neurons react to the faces of specific people. The Jennifer Aniston cell, also known as a grandmother cell, is a putative type of neuron that switches on in response to the face of a recognizable person—a celebrity or a close relative.

I told Rodrigo I thought our cells could be the building blocks for his cells, without thinking very deeply about how this would work. That night, sleepless from jet lag, I recognized a major difference between our face cells and his. I had described in my talk how our face cells computed their response to weighted sums of different face features. In the middle of the night, I realized this computation is the same as a mathematical operation known as the dot product, whose geometric representation is the projection of a vector onto an axis (like the sun projecting the shadow of a flagpole onto the ground).

Remembering my high school linear algebra, I realized this implied that we should be able to construct a large "null space" of faces for each cell—a series of faces of varying identity that lie on an axis perpendicular to the axis of projection. Moreover, all these faces

DORIS Y. TSAO (*face images*)

© 2019 Scientific American

would cause the cell to fire in exactly the same way.

And this, in turn, would suggest cells in face patches are fundamentally different from grandmother cells. It would demolish the vague intuition everyone shared about face cells—that they should be tuned to specific faces.

I was the first person in the meeting's breakfast hall at 5 A.M. the next morning and hoped to find Rodrigo so I could tell him about this counterintuitive prediction. Amazingly, when he finally showed up, he told me he had the exact same idea. So we made a bet, and Rodrigo allowed the terms to be framed in a way that would be win-win for me. If each cell really turned out to have the same response to different faces, then I would send Rodrigo an expensive bottle of wine. If, on the other hand, the prediction did not pan out, he would send me solace wine.

In search of an answer back in our lab at Caltech, Le Chang first mapped the preferred axis for a given cell using responses to the 2,000 faces. Then he generated, while still recording from the same cell, a range of faces that could all be placed on an axis perpendicular to the cell's preferred axis. Remarkably, all these faces elicited exactly the same response in the cell. The next week Rodrigo received an exquisite bottle of Cabernet.

The finding proved that face cells are not encoding the identities of specific individuals in the IT cortex. Instead they are performing an axis projection, a much more abstract computation.

An analogy can be made to color. Colors can be coded by specific names, such as periwinkle, celandine and azure. Alternatively, one can code colors by particular combinations of three simple numbers that represent the amount of red, green and blue that make up that color. In the latter scheme, a color cell performing a projection onto the red axis would fire electrical impulses, or spikes, proportional to the amount of red in any color. Such a cell would fire at the same intensity for a brown or yellow color containing the same amount of red mixed in with other colors. Face cells use the same scheme, but instead of just three axes, there are 50. And instead of each axis coding the amount of red, green or blue, each axis codes the amount of deviation of the shape or appearance of any given face from an average face.

It would seem then that the Jennifer Aniston cells do not exist, at least not in the IT cortex. But single neurons responding selectively to specific familiar individuals may still be at work in a part of the brain that processes the output of face cells. Memory storage regions—the hippocampus and surrounding areas—may contain cells that help a person recognize someone from past experience, akin to the famed grandmother cells.

Facial recognition in the IT cortex thus rests on a set of about 50 numbers in total that represent the measurement of a face along a set of axes. And the discovery of this extremely simple code for face identity has major implications for our understanding of visual object representation. It is possible that all of the IT cortex might be organized along the same principles governing the face-patch system, with clusters of neurons encoding different sets of axes to represent an object. We are now conducting experiments to test this idea.

NEURAL ROSETTA STONE

IF YOU EVER GO to the British Museum, you will see an amazing artifact, the Rosetta stone, on which the same decree of Memphis is engraved in three different languages: Egyptian hieroglyphics, Demotic and Ancient Greek. Because philologists knew Ancient Greek, they could use the Rosetta stone to help decipher Egyptian hieroglyphics and Demotic. Similarly, faces, face patches and the IT cortex form a neural Rosetta stone—one that is still being deciphered. By showing pictures of faces to monkeys, we discovered face patches and learned how cells within these patches detect and identify faces. In turn, understanding coding principles in the face-patch network may one day lead to insight into the organization of the entire IT cortex, revealing the secret to how object identity more generally is encoded. Perhaps the IT cortex contains additional networks specialized for processing other types of objects—a whirring factory with multiple assembly lines.

I also hope that knowing the code for facial identity can help fulfill my college dream of discovering how we imagine curves. Now that we understand face patches, we can begin to train animals to imagine faces and explore how neural activity is shaped by the purely internal act of imagination. Lots of new questions arise. Does imagination reactivate the code for the imagined face in the face patches? Does it bring back even earlier representations of contours and shading that provide inputs to the face-patch system? We now have the tools to probe these questions and better understand how the brain sees objects, imagined or real.

Because almost all the brain's core behaviors—consciousness, visual memory, decision-making, language—require object interactions, a deep understanding of object perception will help us gain insight into the entire brain, not just the visual cortex. We are only starting to solve the enigma of the face. SA

Doris Y. Tsao is a professor of biology at the California Institute of Technology and an investigator of the Howard Hughes Medical Institute. She is also director of the Tianqiao and Chrissy Chen Center for Systems Neuroscience at Caltech. In 2018 she was named a MacArthur Fellow.

MORE TO EXPLORE

The Code for Facial Identity in the Primate Brain. Le Chang and Doris Y. Tsao in *Cell*, Vol. 169, No. 6, pages 1013–1028; June 1, 2017.
How Do We Recognize a Face? Rodrigo Quian Quiroga in *Cell*, Vol. 169, No. 6, pages 975–977; June 1, 2017.

scientificamerican.com/magazine/sa

© 2019 Scientific American

© 2019 Scientific American

THE POINT OF A PERIOD

Age-old taboos against menstruation have led to a lack of research on how women's cycles work, with serious consequences for their health

By Virginia Sole-Smith

Photographs by Jamie Chung

© 2019 Scientific American

IN 2007 SUSAN BROWN ENCOUNTERED THE REPELLING POWER OF PERIOD BLOOD. While studying what menstrual fluid might reveal about a woman's health, she wanted data from a cross section of subjects beyond the student volunteers at the University of Hawaii at Hilo, where she worked as an evolutionary psychologist. Brown's team members set up a booth near the entrance of a Walmart in downtown Hilo and hung a sign that said, "Menstrual Cycle Research." Then they waited. All afternoon women and men would spot the sign, then gingerly skirt past without making eye contact.

About six months later Brown and her Hilo colleague Lynn Morrison presented their findings at the annual meeting of the American Association of Physical Anthropologists. A wave of "nervous twittering" broke out when Morrison described carrying menstrual blood samples down the hallway of their laboratory to analyze hormone levels and other biomarkers. "The audience was fine discussing a woman's cycle in the abstract," Brown explains, "but not menstrual blood itself."

That aversion has influenced women's relationships to their own bodies as well as how the medical establishment manages women when things go wrong with their reproductive health. "Our menstrual taboo is at the core of how this science is getting done," Brown says of research on menstruation.

Or not getting done, as the case may be. It is hard to measure how much money is spent on period research, but experts agree the subject is underfunded. "It's a chicken-and-egg situation, where there's not much funding for research, so there's also not much quantifying of that lack of research," says Elizabeth Yuko, a bioethicist at Fordham University.

Yet period disorders are incredibly common. When Saudi Arabian researchers surveyed 738 female college students in a 2018 study, they found that 91 percent reported at least one menstrual problem: some got their periods irregularly or not at all; others reported excessive levels of bleeding and pain. Different studies show that as many as one in five women experiences menstrual cramps severe enough to limit her daily life. About one in 16 worldwide suffers from endometriosis, a disease where menstrual blood and tissue migrate outside a woman's uterus and form painful lesions in her pelvic cavity. And one in 10 women has polycystic ovarian syndrome, a hormonal imbalance that disrupts a woman's cycle and is a leading cause of infertility. "You can argue we need to put our resources toward researching the life-and-death stuff," Yuko says. "But that argument falls apart because we've had no problem funding erectile dysfunction research."

Menstruation, of course, is essential to human reproduction and therefore survival. It is also one of the biological processes that makes us special because humans, chimpanzees, bats and elephant shrews are among the only animals on earth that go through it. The vast majority of mammals signal fertility through estrus, the period when females are ovulating and display their sexual receptivity via genital swelling, behavioral changes or pronounced alterations in body odor. The female human body, however, conceals this critical window. Instead our most visible sign of potential fertility is menstrual blood, which, ironically, appears after the fertile period has closed. The endometrial lining of the uterus thickens over the course of a woman's cycle as her estrogen level rises. If none of the eggs she releases at ovulation joins with a sperm and implants in that lining as a fertilized zygote, then levels of estrogen and another hormone called progesterone drop, triggering the uterus to shed the thickened endometrium so it can start fresh in the next cycle.

But beyond this basic picture, scientists are still struggling to understand fairly fundamental questions: Why do we share this process with at least six species of bats, for example, but not monkeys? And just what is menstrual blood, exactly? "It's quite different from regular blood," Brown notes. "We know it can't clot and is full of immune agents, but we don't know much about what they do." It is also unclear why we shed this biological tissue so dramatically when most mammals that experience estrus appear to reabsorb their endometrial linings at the end of each cycle. Even less is known about why so many women—up to 80 percent by some estimates—experience cramps, bloating, fatigue, anger or other symptoms just before the onset of menstruation. "We know

IN BRIEF

Squeamishness about female menstruation has led to limited research on how periods work and why they evolved.

Yet many women experience serious pain and other problems with their periods, and medicine has few treatments.

Scientists are just beginning to grapple with questions such as whether it is healthy to have many periods throughout a lifetime and whether suppressing menstruation through birth control is beneficial.

© 2019 Scientific American

so little about menstruation," says Tomi-Ann Roberts, president of the Society for Menstrual Cycle Research and a professor of psychology at Colorado College, and what scientists do know is often badly communicated with the public. "Because of this, our attitudes toward menstruation are overwhelmingly negative. This has real consequences for how we can begin to understand healthy menstruation, as well as menstruation-related disorders and the treatment options available."

MENSTRUAL CUPS can be a reusable, environmentally friendly alternative to tampons and pads.

MASKING MENSTRUATION

THE TABOO has taken many forms. In 1920 a Hungarian-born pediatrician working in Vienna named Béla Schick published a collection of anecdotal observations: When he asked a menstruating woman to handle flowers, they wilted within minutes. When he compared the bread dough made by several women, the loaf made by the one having her period rose 22 percent less. Schick concluded that menstrual blood contained a kind of poison. By the early 1950s Harvard University scientists were referring to "menotoxins" and injecting menstrual blood into animals to observe the effects. Some of those animals died, most likely because the blood samples carried bacteria and other contaminants. Not much came of these experiments in terms of useful data, but the notion that menstrual blood contains mysterious and even dangerous properties has persisted in the scientific literature and our cultural imagination.

By the late 1950s research around menstruation had shifted to center almost entirely on preventing unplanned pregnancies at a time when maternal and infant mortality was troublingly high, especially in poor communities. In 1923 Margaret Sanger, the activist, nurse and founder of the organizations that would later become Planned Parenthood, wrote that "Birth Control means liberation for women and for men." In 1951 she met a physiologist named Gregory Pincus, who had performed what was considered at the time to be the first in vitro fertilization of rabbits. With Sanger securing funding, Pincus set up a lab to test formulations of synthetic versions of hormones that regulate the menstrual cycle and teamed up with John Rock, a Boston obstetrician-gynecologist, to run clinical trials of the drug.

After a study of almost 60 women in and around Boston, Pincus and Rock turned to Puerto Rico to run the first large-scale trial of the drug that the U.S. Food and Drug Administration would approve in 1960 as the first oral contraceptive. They recruited 265 Puerto Rican women, many of them poor, to the study without the level of "informed consent" required today. Twenty-two percent of the participants dropped out after reporting side effects such as nausea, dizziness, headaches and vomiting. The study's medical director argued that the pill "caused too many side reactions to be generally acceptable." Nevertheless, it went to market.

The pill was, of course, celebrated as a huge breakthrough. "It was the first form of birth control separate from sex that women could completely control," notes Elizabeth Kissling, a professor of women and gender studies at Eastern Washington University. It is impossible to overstate the freedom the pill represented for women, whose reproductive lives were otherwise largely under male control. But liberation came with a price. By the late 1960s patients across the U.S. were reporting the same symptoms documented during the Puerto Rican trial. Despite many reformulations over the ensuing decades, side effects remain a problem for many women on the pill; risks for breast cancer, blood clots and stroke may also be higher. In their quest to bring reproductive freedom to women, Sanger, Pincus and Rock appear to have ignored the implications of shutting down a woman's natural cycle, Kissling explains. In other words, scientists figured out how to supplant periods long before they began trying to understand why they work the way they do.

It was not until the late 1980s that scientists really began to grapple with the larger question of why menstruation happens at all. As an undergraduate, evolutionary biologist Beverly I. Strassmann wrote a paper on how concealing ovulation could entice more paternal partners. (Because a woman's fertile window is more or less invisible, it encourages what researchers call pair-bonding: human males invest in fewer sexual relationships and protect and care for the resulting offspring as a way to ensure their paternity.) Strassmann, now a professor of anthropology at the University of Michigan, wanted to explore human attitudes toward menstruation by collecting data in a community where women spend five nights of their period sleeping in huts that are separate from the rest of the tribe.

In 1986 Strassmann moved to Mali to conduct field research on the Dogon, an ethnic group of millet farmers that hew to their traditions. Dogon people who continue to practice their indigenous religion believe that a menstruating woman's presence would desecrate the religious objects in the family compounds. Researchers had not previously considered that these religious beliefs were rooted in any kind of reproductive agenda. But, as Strassmann explains, she hypothesized that this was "a

© 2019 Scientific American

cultural pattern embedded in religion that did directly serve reproduction." Although research on modern indigenous communities can offer only clues about how humans lived thousands of years ago, Strassmann hoped to show that long-standing cultural taboos around menstruation had developed to support our larger evolutionary goals.

During her initial fieldwork, Strassmann studied the community's use of menstrual huts for almost three years, collecting urine samples from 93 women to test hormone levels and prove that their use of the huts correlated with actual menstruation patterns. She also observed how quickly most of the women got pregnant again after their visits to the huts. Although the practice was ostensibly about keeping menstruation sequestered, the huts themselves were located in full view of a shade shelter used by men in the community. So the huts made a woman's fertility status clear to her husband and his family whether she liked it or not. (As noted earlier, women enter their "fertile window" after their period.)

Other religious practices around menstruation, such as the Orthodox Jewish purification ritual of sending menstruating women to mikvah baths, can also be traced to men's need to track female fertility and schedule sexual activity accordingly. And although the advent of the pill means that many women can now control their reproductive life in ways that render the purpose of such practices moot, the taboos still persist, Roberts says. "We still think of menstruation as something that women have to keep hidden and separate."

PERIOD EVOLUTION

ALTHOUGH STRASSMANN'S WORK was primarily about understanding the biological underpinnings of menstrual taboos, her data also revealed important characteristics about the process of menstruation itself. Perhaps her most oft-cited finding was published in 1997 in *Current Anthropology:* across human history, menstruation has been a rather infrequent event. That is because women tend to get pregnant earlier, have more babies and spend more time breastfeeding in communities where birth control is unavailable or difficult to access than they do in communities with high rates of birth-control usage. "We think of periods as happening 12 times a year, but if you're pregnant and then nursing for extended time frames, that's a stretch of two or three years for each child when you're not menstruating," Strassmann explains. Her data showed that in the 1980s the average Dogon woman menstruated only around 100 times in her life, compared with the average American woman's experience of as many as 400 periods in her lifetime. And Dogon women's experience is closer to what all women would have experienced throughout history before the development of the pill.

This historical infrequency of menstruation helps to explain why humans evolved to do something as potentially disadvantageous as releasing blood—losing iron, protein and other nutrients and probably attracting predators in the process. It could also help explain why periods and the week before their onset can be so unpleasant for many women. Michael Gillings, a professor of molecular evolution at Macquarie University in Australia, became interested in women's experiences of premenstrual symptoms (PMS) when premenstrual dysphoric disorder (PMDD) was added to the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* in 2013.

PMDD is defined as severe irritability, depression or anxiety in the week or two prior to menstruation, with symptoms easing two or three days after menstruation begins. But Gillings, along with many feminist scholars, balked at the characterization of mood swings as disordered. "Up to 80 percent of women report these symptoms; that makes PMS normal, not a psychological disorder," he says. "So we have to ask, Was there, at some point in history, an advantage to having these symptoms?" In 2014 he published a paper in the journal *Evolutionary Applications* arguing that PMS offered a selective advantage because it caused tension between pair-bonds and therefore might help women dissolve relationships with infertile men. "It is difficult to prove a hypothesis like this," he acknowledges. And the media response characterized him as insensitive to the suffering of women. "I was burned in effigy on five continents," he says. Some researchers counter Gillings's claim that PMS is a product of evolution and contend that its roots are more cultural than biological because it manifests differently around the world. Roberts sees the concept mostly as one influenced by the menstrual taboo and a way to dismiss women's emotions.

Scientists are also divided over whether the act of bleeding itself serves an evolutionary purpose. "It's never made sense to me that we have this free-flowing blood, while other animals reabsorb it," Brown argues. Many evolutionary biologists now think that the essential feature of women's cycles is not the bleeding but rather the ability of the uterus to thicken its lining in preparation for implantation and then dispose of the endometrium when it is not needed. "A healthy endometrium requires constant metabolic support, so it is less energy-intensive for the female body to tear down and rebuild it each cycle than it is to maintain it in a constant state of readiness for embryo implantation," Strassmann explains. Human circulation happens to result in a particularly bloody endometrium. "Our physiology does not permit reabsorption, so much of the blood gets discharged as menstruation," she says. Bleeding may therefore be an insignificant by-product of evolution rather than an advantage.

A WORLD WITHOUT PERIODS?

IF THE ACT OF SHEDDING menstrual blood poses no clear health benefit or evolutionary advantage and if, historically, women have not even done it all that often, then why, in this postpill era, do women continue to do it all? The answer: some do not. In January the Royal College of Obstetricians and Gynecologists in London released new guidelines that approved skipping the placebo pills in birth control to reduce the frequency of periods or avoid them altogether.

Although this formal acknowledgment is new, the practice is not. Medical menstrual suppression has long been embraced by clinicians, the media and women frustrated by the pain, mood swings or inconvenience of their menstrual cycle. The pharmaceutical industry also took notice: as the researcher who first measured and quantified the frequency of human menstruation, Strassmann has been asked to present her data to drug manufacturers, who have offered several versions of the pill and other forms of contraception that are formulated to let women skip their periods more often, if not avoid them altogether.

Skipping that monthly ordeal can mean avoiding debilitating

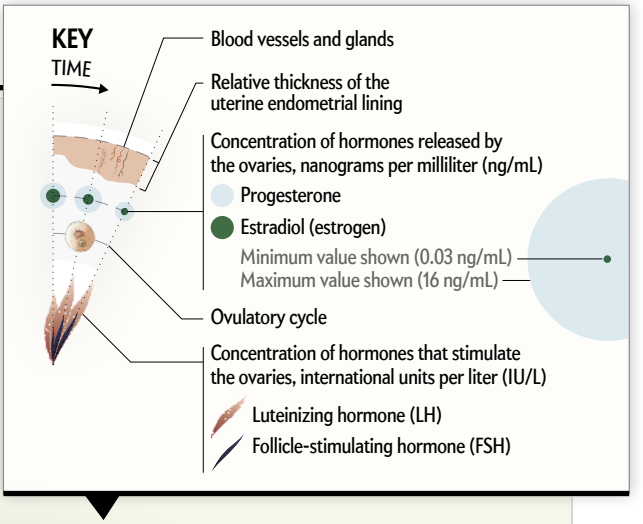
Continued on page 19

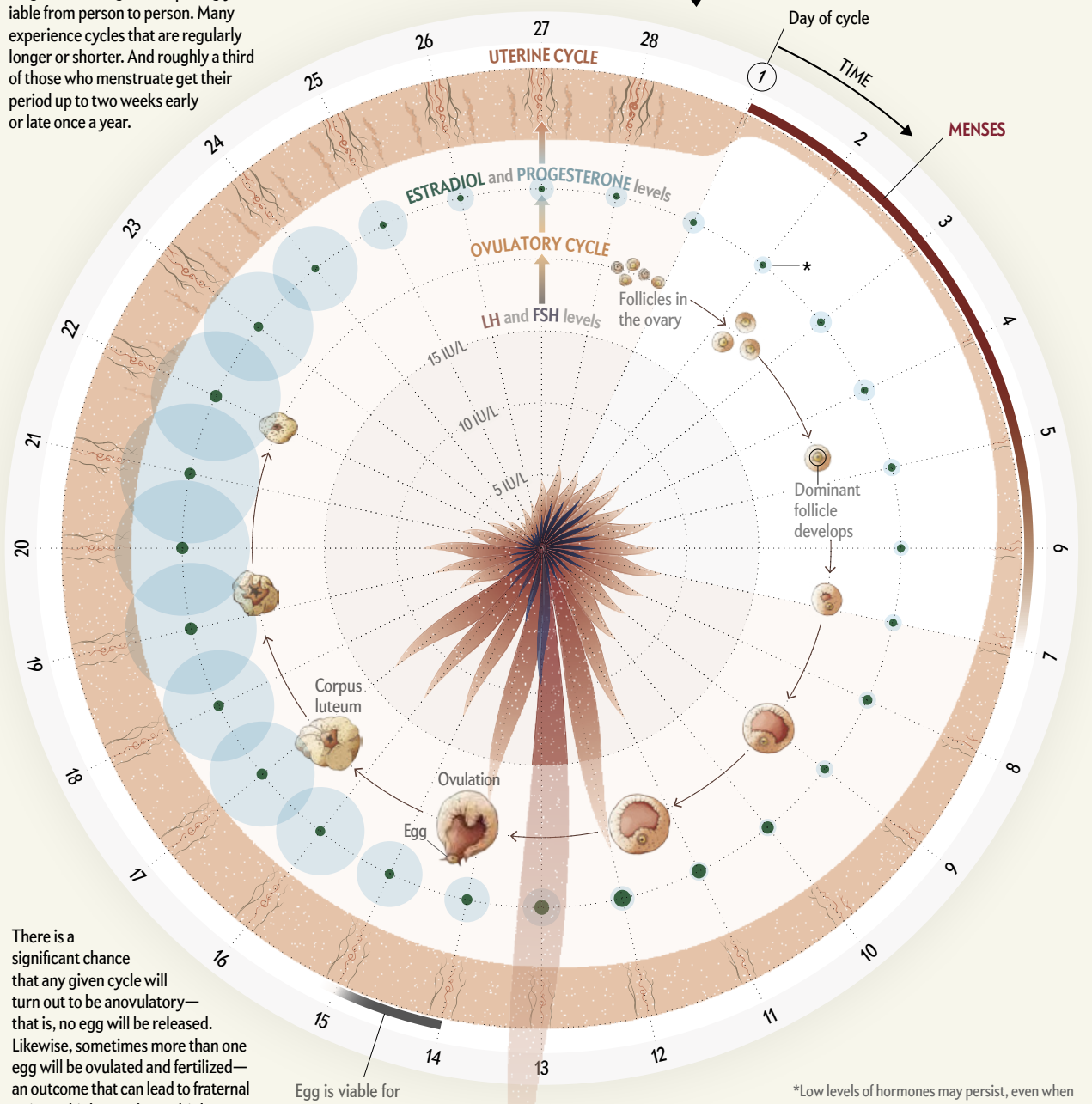
© 2019 Scientific American

The Menstrual Cycle

Humans are among the very few species to experience a period. The menstrual cycle starts in the brain, which sends signals to the pituitary gland (*not shown*) to produce hormones that stimulate the ovaries. The ovaries house egg-containing follicles that release an egg during ovulation. The ovaries also secrete hormones to help prepare the uterus to host an embryo, which results if the egg is fertilized by a sperm. If no embryo implants, the uterus disposes of its lining, and the cycle begins again.

The average menstrual cycle is 28 days long, but the length is surprisingly variable from person to person. Many experience cycles that are regularly longer or shorter. And roughly a third of those who menstruate get their period up to two weeks early or late once a year.


KEY
TIME
Blood vessels and glands
Relative thickness of the uterine endometrial lining
Concentration of hormones released by the ovaries, nanograms per milliliter (ng/mL)
Progesterone
Estradiol (estrogen)
Minimum value shown (0.03 ng/mL)
Maximum value shown (16 ng/mL)
Ovulatory cycle
Concentration of hormones that stimulate the ovaries, international units per liter (IU/L)
Luteinizing hormone (LH)
Follicle-stimulating hormone (FSH)



There is a significant chance that any given cycle will turn out to be anovulatory—that is, no egg will be released. Likewise, sometimes more than one egg will be ovulated and fertilized—an outcome that can lead to fraternal twins or higher-order multiples.

SOURCE: "CIRCULATING IMMUNOREACTIVE INHIBIN LEVELS DURING THE NORMAL HUMAN MENSTRUAL CYCLE" BY R. I. McLACHLAN ET AL., IN *JOURNAL OF CLINICAL ENDOCRINOLOGY AND METABOLISM*, VOL 65, NO. 5; NOVEMBER 1987 (*hormone level data*); JERILYNN C. PRIOR *University of British Columbia, Center for Menstrual Cycle and Ovulation Research (consultant)*

Graphic by Federica Fragapane

© 2019 Scientific American

Why Are Girls Getting Their Periods So Young?

Female puberty is starting earlier and earlier, with worrying consequences for women's health

Three weeks before her eighth birthday, Josie got her period at school. Magen, her mother, stopped at a drugstore for supplies before picking up her daughter. In the tampon aisle, she found a shelf of "tween" menstrual pads promising to "fit smaller bodies." She remembers thinking, "How does this even exist as a product?"

Magen was heartbroken that her seven-year-old was menstruating but not completely surprised. She had begun to notice her daughter's body odor when Josie was six. By the time Josie turned seven, she was getting blackheads on her nose, slamming doors and sleeping late. She developed breast buds the summer before second grade. "That was traumatizing for both of us," Magen says.

Magen showed Josie how to put a pad in her underwear and then called the pediatrician, expecting to be referred for some kind of hormonal testing. Instead, Magen recalls, "he said, 'Yes, this happens. She likely won't be regular for a while, but she's very clearly in puberty at eight years old.'"

The average age of menarche, or a girl's first period, in the U.S. is now 12, according to the most recent data from the Centers for Disease Control and Prevention's National Health and Nutrition Examination Survey, down from 14 a century ago and as much as six months earlier than 20 or 30 years ago. But puberty does not start with menstruation. The onset of breast development, or thelarche, tends to come first, just as Josie experienced. "We're now seeing thelarche occur 18 months to two years earlier than we did a few decades ago," says Frank Biro, who studies problems related to pubertal maturation at Cincinnati Children's Hospital Medical Center. His research, published in 2013 in the journal *Pediatrics,* put the average age of breast development at 8.8 years old for African-American girls, 9.3 for Hispanic girls, 9.7 for Caucasians and 9.7 for Asian-Americans. "The age of breast development has clearly dropped, while the age of menarche has drifted down. They are both concerning," he says.

One popular misconception about menarche is that it represents the onset of ovulation; in fact, most girls do not begin ovulating regularly for up to two years after their first period, which is why early ones can be light and irregular. Menarche is instead triggered by changes in a girl's estrogen levels. The most probable explanation for why periods and breast development might be happening younger is that girls tend to weigh more today than they did a generation ago—and this higher body fat percentage is leading to earlier activation of the pituitary gland, which produces the hormones responsible for puberty.

In Biro's study, a higher body mass index (BMI) was the strongest predictor of early breast development across all racial groups, although the relation was correlative, not causal. "What we need to ask is, Why has BMI gone up?" he says. "Decreased physical activity and a more calorically dense diet are probably part of the puzzle. But I think another critical piece is our ubiquitous environmental exposure to endocrine-disrupting chemicals," or EDCs. This class of chemicals (including phthalates, bisphenol A and others) is used in many consumer products (shower curtains, plastic bottles, couch cushions) and has been shown to mimic estrogen and other naturally occurring hormones in the human body. Biro theorizes that some of these chemicals may promote weight gain or contribute to early puberty by influencing how cells and the body regulate metabolism, which then affects estrogen production. He is currently leading a study tracking the growth and development of 379 girls from age six onward that has been examining relations between their pubertal development and environmental exposures. Trauma could be another explanation: "Stress can also change your estrogen levels," Biro says.

Josie's mom found a shelf of "tween" menstrual pads promising to "fit smaller bodies." She remembers thinking, "How does this even exist as a product?"

To Magen, the more pressing question is not why Josie started puberty so early but rather what this means for her daughter's immediate and long-term health. The data Biro is collecting now show that girls who start puberty early tend to stay in the stage longer, meaning they spend more time in a "window of susceptibility"—a time when the human body is in a particularly critical stage of development, such that environmental exposures and other experiences are more likely to have an impact on their later health. When it comes to future risk of breast cancer, for example, fetal development and infancy are one window, and puberty is the other. "We know that for every year you delay menarche, you decrease the risk of premenopausal and postmenopausal breast cancer by 4 to 8 percent," Biro says. "On a population basis, that's really important."

Other researchers are looking at how early puberty affects girls socially and emotionally. "We know that early reproductive development is not matched by early cognitive development," says Marcia Herman-Giddens, an adjunct professor of maternal and child health at the University of North Carolina's Gillings School of Global Public Health. "So how do we teach children to manage sexual urges and other realities of puberty? And of course, these girls have to deal with sexual advances from older boys and even men long before they are ready to navigate that."

Magen is trying to figure out how to introduce these issues to Josie in an age-appropriate way without overburdening her already anxious daughter. "I've had to tell her, 'At some point, you're going to feel interested in relationships and sex, and when you are, you need to tell me right away,'" she says. "But am I really going to have to put a 12-year-old on birth control to make sure she's safe?" —*V.S.-S.*

© 2019 Scientific American

Continued from page 16

pain, prolonged heavy bleeding, migraines and other symptoms that can dramatically impair a woman's quality of life. The approximately 25 percent of reproductive-age women and girls who struggle with additional kinds of severe menstrual pain may be at increased risk for developing other chronic pain conditions. "We suspect the cyclical experience of monthly menstrual pain somehow alters how some women process all kinds of pain," explains Laura Payne, who directs pain research for a pediatric program at the David Geffen School of Medicine at the University of California, Los Angeles.

To many doctors faced with patients whose periods cause problems, "the pill is the closest thing we have to a panacea in women's health," says Jonathan Schaffir, an associate professor of obstetrics and gynecology at the Ohio State University. But is it? "The pill isn't a treatment for these conditions," Kissling says. "It's a way of refusing to treat them." It can take up to a decade or longer from disease onset for a woman to be diagnosed with endometriosis, for example, in part because doctors are so quick to prescribe the drug to teenagers reporting bad cramps without investigating to see if there is an underlying cause, says endocrinologist Jerilynn C. Prior of the University of British Columbia. And where one version of the pill may succeed in masking a woman's symptoms, another may exacerbate them. "You can spend years jumping from one pill to another, not finding relief," notes Kissling, who published a paper on how women end up "treating each other," for better or worse, in online forums, where they share alternative medicine remedies and other tips out of frustration with their doctors' limited repertoire.

Strassmann and many others are skeptical about the health effects of medically induced menstrual suppression, which may expose women to hormone levels higher than what they would have experienced in the evolutionary past or even now, when regularly cycling on the pill. "It's true a monthly menstrual period is not necessary," she says. "But taking more progestin to skip your period is not living like our ancestors did 500 or 1,000 years ago." Research shows taking the pill reduces the risk for endometrial and ovarian cancers but slightly raises the risk for breast cancer, stroke and blood clots.

In 2017 Strassmann and her colleagues published a paper in *Evolution, Medicine, and Public Health* tracking how exposure to synthetic hormones varied depending on the type of birth-control pills used. "We know that American women experience more periods than the Dogon because they start menstruating earlier and have fewer children, and we know that having more periods is associated with a higher breast cancer risk," she explains, noting that the relation is likely because of the additional hormone exposures accrued from those extra periods. "But we don't really know how that risk squares with the hormone exposure women are also getting from long-term use of birth-control pills." After analyzing data from 12 studies, as well as the information on birth-control package inserts, Strassmann's team concluded that some types of the pill exposed women to a quadruple dose of progestin (a synthetic form of progesterone contained in the pill), relative to the progesterone their naturally cycling body would produce.

PHONE APPS, such as Clue (*shown*), help women to keep track of their monthly cycles.

Nobody knows for sure what that exposure to synthetic hormones will mean long term for women using the pill to suppress their cycles indefinitely. This knowledge gap speaks to broader concerns about our ignorance around menstruation. If Rock and Pincus had begun their work with a deeper understanding of menstruation's evolution and purpose, how might that have affected the pill's development? Would women today have more—and more targeted—options to manage their menstrual pain and associated disorders?

In this latest iteration of our menstrual taboo, dispatching with the period instead of researching its complexity might have unforeseen health consequences, Prior says. "Our data on the pill come from generations of women who followed the schedule for 28-day cycles and didn't stay on it for nearly as long as women do today," Kissling says. "What we have now" with women using birth control for long-term suppression "is the largest uncontrolled medical experiment on women in history." SA

Virginia Sole-Smith is a journalist who writes about feminism, body image and health for the *New York Times Magazine, Harper's Magazine, Elle* and many other publications. She is author of *The Eating Instinct: Food Culture, Body Image, and Guilt in America* (Henry Holt, 2018).

MORE TO EXPLORE

The Biology of Menstruation in *Homo sapiens*: Total Lifetime Menses, Fecundity, and Nonsynchrony in a Natural-Fertility Population. Beverly I. Strassmann in *Current Anthropology,* Vol. 38, No. 1, pages 123–129; February 1997.

FSRH Guideline: Combined Hormonal Contraception. Faculty of Sexual and Reproductive Healthcare of the Royal College of Obstetricians and Gynecologists, January 2019 (amended July 2019). www.fsrh.org/standards-and-guidance/documents/combined-hormonal-contraception

scientificamerican.com/magazine/sa

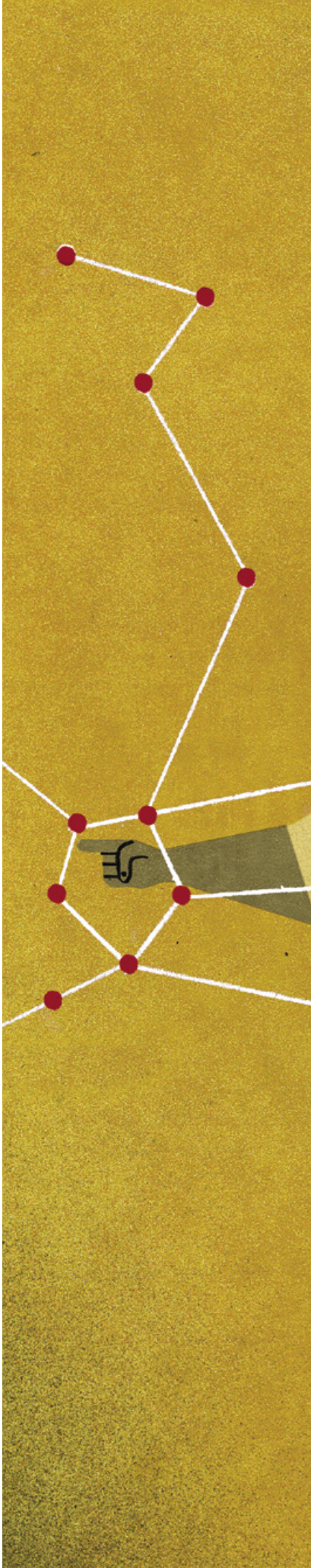
© 2019 Scientific American

Vaccines Reimagined

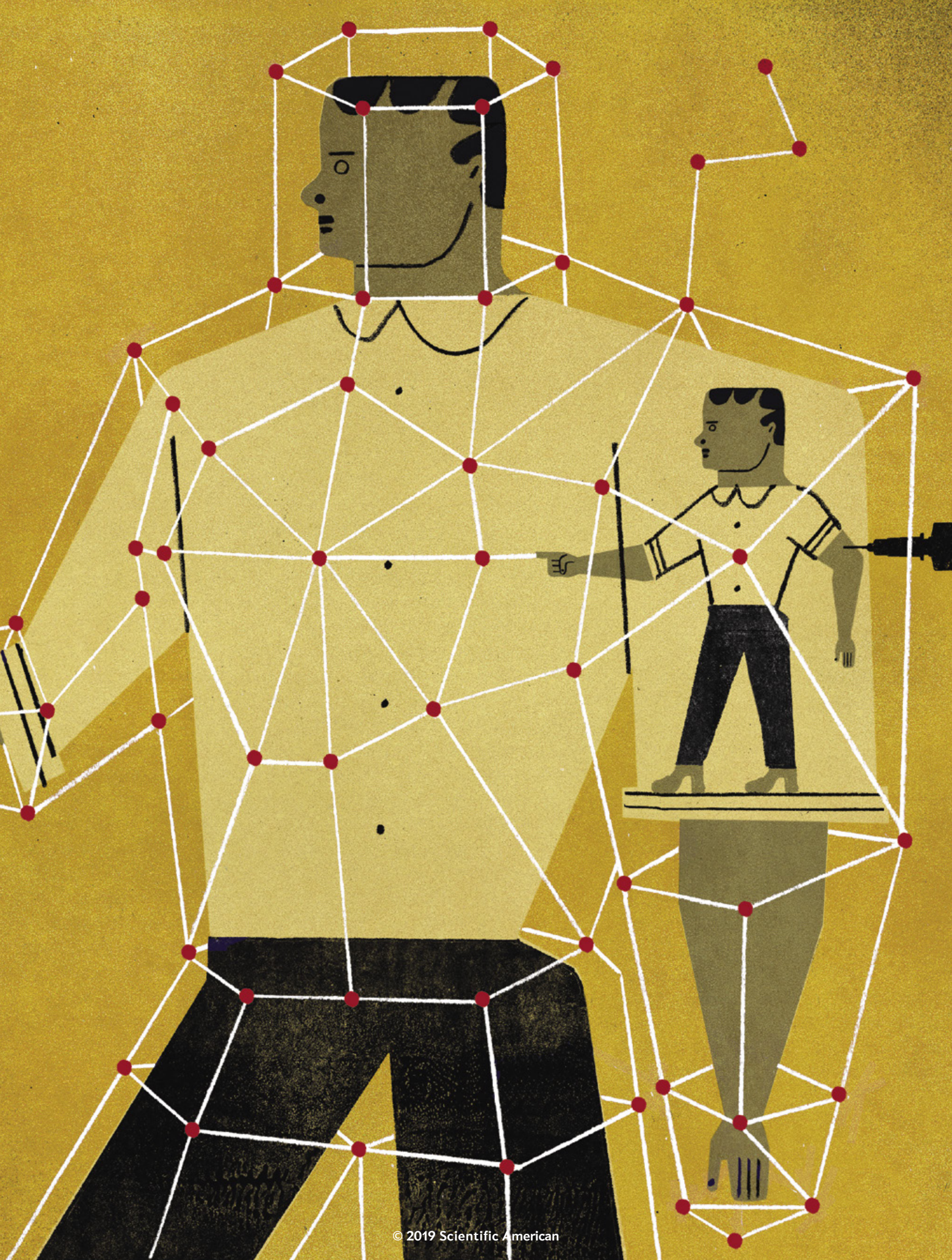
A controversial theory holds that one immunization, given properly, can protect against many diseases besides its target

By Melinda Wenner Moyer

Illustration by David Plunkert



© 2019 Scientific American



© 2019 Scientific American

THE HEAT OF THE SUN, A BLAZING BASKETBALL IN THE WEST AFRICAN SKY, was softened by a breeze one afternoon last spring. Every so often the wind whisked a mango off a tree branch and dropped it with a thud on the corrugated iron roof that covered the health center in Bissau, the biggest city in the tiny country of Guinea-Bissau, where the rust-colored ground hadn't felt a raindrop in six months. Inside the building, the air was still and dry, and a line of women and toddlers were sticky with sweat.

An 18-month-old named Maria with thick, dark braids studied me nervously as she perched on her mother's lap. (The child's name has been changed to protect her privacy.) Next to them, Carlito Balé, a soft-spoken doctor in a short-sleeved, white button-down shirt, talked with Maria's mother in Portuguese creole, a percussive fusion of Portuguese and African dialects. Balé was telling the mother that Maria was eligible to participate in a clinical trial to test whether an extra dose of measles vaccine prevented not just the measles but many childhood infections that cause serious illness and death.

IN BRIEF

Vaccines target specific diseases, but a line of studies suggests that some offer much broader protection. **Live immunizations** in particular may cut child mortality rates by 50 percent overall, research indicates. **This work,** spearheaded by Peter Aaby and Christine Benn in Guinea-Bissau, has also drawn criticism for overstated conclusions.

In the U.S., where life-threatening infections are rare, such a trial might not garner many volunteers. But in Guinea-Bissau, where lives have been scarred by decades of scant resources and poor medical care, families lined up in droves. The nation is one of the world's poorest, and the CIA ranks infant mortality there as the fourth highest among 225 countries. Mothers often wait months to name their babies because one out of every 12 will die before his or her first birthday.

The researchers leading the trial—anthropologist Peter Aaby and physician Christine Benn, whom I had traveled to Guinea-Bissau to meet—have amassed evidence that a few specific vaccines can thwart a multitude of threatening plagues. Over decades they have published hundreds of studies suggesting that live, attenuated vaccines, which are made from weakened but living viruses or bacteria, can stave off not just their target infections but other diseases, such as respiratory infections (including pneumonia), blood infections (including sepsis) and diarrheal infections. In a 2016 review published in the journal *BMJ,* a research team commissioned by the World Health Organization analyzed 68 papers on the topic, many of which came from Aaby and Benn's research. It concluded that the measles and tuberculosis vaccines "reduce overall mortality by more than would be expected through their effects on the diseases they prevent." Some of the research the team evaluated linked the measles vaccine with a whopping 50 percent lower risk of death from any cause.

This notion that live vaccines have what are called "off-target" effects—and powerful ones—has implications that stretch far beyond Africa. In 2017 in the U.S., for instance, researchers at the Centers for Disease Control and Prevention reported that children were half as likely to be hospitalized for nonvaccine-targeted infections between the ages of 16 and 24 months if the last immunization they had received was a live vaccine rather than an inactivated one. New research in immunology suggests that live vaccines can have such wide-ranging effects because they stimulate a part of the immune system that fights a broad-based war against all outside invaders, giving the system a head start on defense. "Although we still need to know much more about the details, I now have no doubt that vaccines do have some off-target effects because of the support from many different types of evidence," says Frank Shann, a pediatrician at Royal Children's Hospital Melbourne in Australia.

Yet other scientists are far less certain. Aaby and Benn's work is, in fact, quite controversial. For one thing, most of the studies from the two Danish researchers do not prove cause-to-effect connections. "Purported effects" is how Paul Fine, an infectious disease epidemiologist at the London School of Hygiene & Tropical Medicine, describes them. Kids who get live vaccines might survive longer for reasons that have nothing to do with immunizations: the children in those groups might have been healthier to begin with. To address these concerns, Aaby and Benn are now running intervention tri-

© 2019 Scientific American

als, such as the one Maria was being recruited for. In it, children will be matched for age and basic health, but some will have only the standard single measles shot at nine months, whereas others will get an additional dose as toddlers.

The two investigators also counter that political and pragmatic concerns drive resistance to their ideas far more than do valid scientific critiques. Aaby says that his and Benn's research is inconvenient for public policy because it indicates that live vaccines should be given last in any vaccine series, which upends current immunization schedules and could inadvertently trigger parental worries about safety. Public health scientists "don't want to hear it, and I can understand why they don't want to hear it," Aaby says. And as a result, he claims, many orthodox vaccine researchers "have clearly made me persona non grata." The 74-year-old, who is bespectacled and has a salt-and-pepper goatee, fits the part of the eccentric, obstinate and misunderstood scientist so well that he has literally become one in a novel: he inspired a character in a best-selling 2013 Danish mystery book, *The Arc of the Swallow,* who gets murdered in the first chapter.

In real life, Aaby and Benn's ideas may be reaching a tipping point. The WHO wrote in a 2014 report that nonspecific vaccine effects seem "plausible and common" and worthy of more attention. Therefore, in April 2017 the agency announced it would oversee the design of two multiyear clinical trials to further test the hypothesis, although those trials have not yet begun. The two researchers, whose professional relationship has evolved into a long-term romantic one, are pushing forward with more of their own trials, too. One of them is the study Maria's mother was considering. As I watched in the health center, she decided to enroll her daughter, so Balé picked up a large envelope containing dozens of smaller sealed envelopes and held it open toward her, telling her to pick one—a step that ensured that her daughter would be randomly allocated to either the treatment or control group. Opening her chosen envelope, Balé announced that Maria would get the extra vaccine, and her mother flashed a hopeful smile. She carried her daughter into the next room, where a nurse in a long, white-and-orange tie-dyed dress, black glasses and a kind smile waited with a needle.

CHRISTINE BENN AND PETER AABY sit in front of their house in Guinea-Bissau. Much of their research on vaccines is conducted in that country.

THE MEASLES CLUE

IN 1979, SOON AFTER launching a health surveillance project in Bissau, a young Aaby watched measles kill one out of every four babies in the area. That was the year he saw his first dead body, and he saw a lot more than one.

Back then, childhood vaccines were rare in Africa. The WHO estimates that in 1980, only 6 percent of African children received the first dose of live measles vaccine, and 8 percent got the first inactivated DTP vaccine, which protects against diphtheria, tetanus and pertussis. It's not as if the vaccine was new; the combination DTP vaccine was licensed in 1949, yet 31 years later fewer than one in 12 African children ever received a dose. Indeed, only a handful of childhood vaccines were even available then in Africa. In addition to the DTP and measles vaccines, there was a live tuberculosis vaccine called bacillus Calmette-Guérin (BCG) and a live polio vaccine. In 1980 in the U.S., on the other hand, 86 percent of kids received the live measles vaccine, 98 percent were inoculated with the inactivated DTP vaccine, and 95 percent had gotten live polio vaccines. African children today receive a lot more vaccines than they used to, but they still woefully lag behind the U.S.

In 1978, a year before the historic measles outbreak began, Aaby had been sent to Guinea-Bissau by a Swedish organization to investigate malnutrition. When the epidemic swept into the city, he pulled strings to import measles vaccines and began to inoculate the local children, all the while keeping track of infection and death rates. The move was a bold one: at that time, public health authorities thought that measles vaccine campaigns in Africa were essentially a waste of money and effort. In a 1981 paper published in the *Lancet,* researchers analyzed survival data after undertaking a measles vaccine campaign in Zaire and concluded that in the future, "it may be useful to think twice before allocating already scarce resources to such a programme." Measles took the lives of the weakest children, they argued; even if the vaccine prevented the infection, the spared children would die from something else soon enough.

Aaby's experience didn't support this argument. The before-and-after numbers he saw were staggering: In 1979, the first year of the outbreak, 13 percent of local children between the ages of six months and three years died; in 1980, when the measles vaccine was available, only 5 percent did. Surprisingly, deaths from causes other than measles dropped by one fifth between 1979 and 1980, too. The trend continued. Even after measles disappeared, immunized children remained more likely than their unvaccinated peers to survive other infections. "It was one of those moments where

SOFIA BUSK HANSEN

© 2019 Scientific American

you can suddenly see something you would never have believed was possible," he recalls. Aaby and his colleagues wrote a letter to the *Lancet* refuting the theory that measles inoculation campaigns in Africa were useless—his first-ever publication in a medical journal. After that, he says of the measles vaccine, "I became obsessed."

Aaby has now published more than 100 studies on this one vaccine. His surveillance program, the Bandim Health Project, a collaboration between Guinea-Bissau's Ministry of Health and the State Serum Institute in Denmark, is one reason why. For more than 40 years the project has been registering all pregnancies, births and deaths in Bissau's urban district of Bandim, as well as in five nearby rural regions. Aaby's team there has monitored the health of more than 500,000 people living in these areas and has collected data on hospitalizations, vaccinations and health-related choices, such as whether people sleep with mosquito nets. One day during my visit, as I walked around with Aaby, a mother holding a baby said she remembered him from when he visited her as a child some 30 years ago. His colleagues and assistants in the project affectionately call Aaby *Homem Grande,* which translates to "Big Man."

Aaby has always been a bit of a lone wolf—he spends many days working by himself in his home office—but less so during the past 15 years. While Benn was in medical school in 1992 at Aarhus University in Denmark, she was advised to reach out to Aaby because she wanted to study whether vitamin A supplementation, routinely given with the measles vaccine in developing countries, interacted in any way with the vaccine. "I still have the piece of paper with his number," Benn, who is 50, tells me as she sits on a bench in Aaby's back garden, her arms hugging her legs. She has been working with Aaby ever since. Benn is now a professor of global health at the University of Southern Denmark and runs the Danish arm of the Bandim Health Project. She is prolific, having published more than 200 papers on issues including the nonspecific effects of vaccines and the impact of vitamin A supplementation on infants in developing countries. She calls Denmark home but spends about 10 weeks a year in Guinea-Bissau. The two researchers bring to the field, and their relationship, complementary personalities: Benn, effervescent and philosophical; Aaby, serious and precise.

For the most part, Aaby and Benn's work on the measles vaccine has supported Aaby's original observations. In a landmark 1995 *BMJ* paper, they analyzed data from 12 previously published studies—some their own—on the association between measles vaccination and mortality in developing countries. They found that the vaccine was linked to a 30 to 86 percent reduction in overall death risk. In each study, measles itself only killed a small proportion of unvaccinated kids, so the vaccine wasn't just preventing measles; something else was going on. In a 2014 paper published in *JAMA,* Aaby and Benn collaborated with Danish researchers to investigate whether these protective effects extended to high-income countries. They found that Danish children who received the live measles-mumps-rubella (MMR) vaccine as their last inoculation were 14 percent less likely to be hospitalized for any infection than were kids who had most recently received the inactivated DTaP-IPV-Hib vaccine for diphtheria, tetanus, acellular pertussis, polio and *Hemophilus influenzae* type B. This study inspired the 2017 analysis by the CDC that found live vaccines to be associated with even stronger protection in the U.S.

Aaby and Benn have also linked the BCG vaccine with lower neonatal mortality, and they have studied the live oral polio vaccine (OPV) as well. In a 2018 paper, they reported that child mortality rates were 19 percent lower after OPV campaigns than before them, and a clinical trial they published in 2015 found that OPV given within two days of birth with BCG reduced mortality risk by 42 percent, compared with BCG alone. Based in part on their findings, 15 scientists wrote a letter to the *Lancet* in 2016 arguing that the global switch from live OPV to IPV, the inactivated polio vaccine, which is part of a plan developed by the international Global Polio Eradication Initiative, could inadvertently increase child mortality.

The two scientists are certain that the evidence they have accumulated points to a clear conclusion: vaccines have more profound effects on the body than we thought. The big mystery they have been grappling with is how, exactly, all this happens.

A BROAD BOOSTER

MIHAI G. NETEA may have an answer. In 2010 Netea, an immunologist at Radboud University in the Netherlands, embarked on a study that he frankly didn't think would be all that interesting. His laboratory was studying how the BCG vaccine affects human immune cells—how it teaches them to recognize and attack the bacterium *Mycobacterium tuberculosis*. To provide an experimental control on one test, lab workers exposed blood samples from vaccinated volunteers to *Candida albicans,* a common yeast. Based on accepted immunology doctrine, which holds that vaccines incite immune responses specific to the targeted pathogen, BCG should have had no effect on the blood's response to *Candida*.

A few weeks later the student running the test approached Netea, concerned. "I think I did something wrong because I see differences with both tuberculosis and with *Candida*," Netea recalls her saying. Perhaps her samples had been contaminated; he suggested that she collect more blood samples and do the experiment over. She did, but the same thing happened. "She came again and said, 'Well, I don't know what to do, but I see precisely the same thing again,'" Netea says. He was flummoxed, so he started reading about BCG and found a handful of surprising animal studies that suggested the vaccine also protected some animals against malaria, influenza and *Listeria monocytogenes,* a common cause of foodborne illness.

That is when Netea's simple study transformed into a Greek siren, a creature beckoning for his full attention. How could a vaccine against tuberculosis change how the body responds to other pathogens? The idea contradicted established paradigms. Immunizations prime the body to make proteins called antibodies that recognize, attach to and attack proteins on the pathogens if the body ever encounters them again. This defense is called adaptive immunity, and it acts like a team of snipers that take out only certain targets. Given adaptive immunity's specificity, it didn't make sense to Netea that it could be responsible for BCG's ability to protect against a number of insults.

Another kind of bodily defense—one that researchers historically thought vaccines had little to do with—is known as innate immunity, and it is more like a battalion told to open fire on anyone who edges into its line of sight. It is the rapid-response team, initiating a fight against any new invader. When pathogens invade, innate inflammatory cells get pulled to the infection site. Large white blood cells called phagocytes—particularly a type called macrophages—engulf and destroy the pathogens. They also secrete immune chemicals called cytokines that draw other immune cells to the scene. The reaction creates proteins

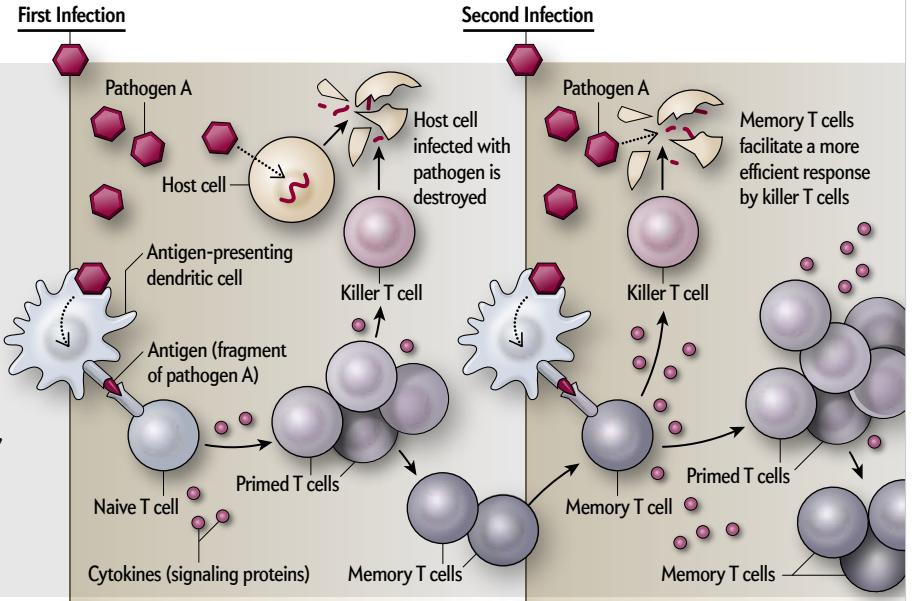
© 2019 Scientific American

Double Defenses

The body's immune system has two arms: adaptive and innate. The adaptive arm creates cells that respond only to specific bacteria or other threats. The innate arm has a faster response, but effectiveness against a particular germ is more limited. A new theory holds that this arm can be "trained" by vaccines with live but weakened pathogens to be more potent against a range of germs.

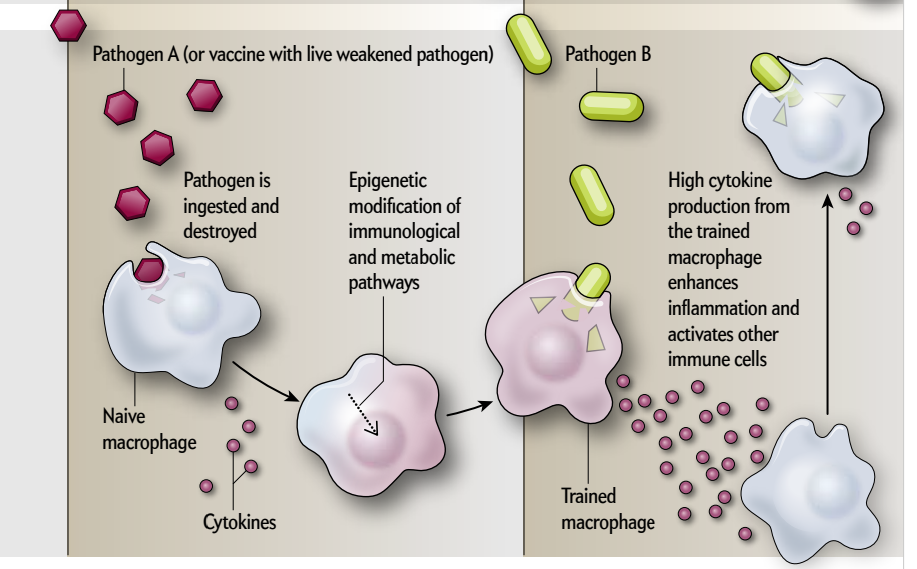
Adaptive Immunity

This part of the immune system begins by capturing pieces of an invading pathogen called antigens. Cells present the antigens—often proteins from bacteria or viruses—to T cells, transforming them from "naive" to "primed." The cells use the antigens to trigger an immune reaction specific to the invader. The response involves killer cells that go after the infected cells, chemical messengers called cytokines that activate other destructive responses and the creation of memory cells that stay in the body to recognize the pathogen, should it show up again. If reinfection happens, memory cells enable the immune system to single out the pathogen and attack it.



Innate Immunity

This arm uses general defense cells called macrophages. They engulf any pathogen and do not have specific targets. But recent research hints that innate components, like adaptive ones, can remember past pathogen encounters. Such encounters may come from a weakened pathogen in a live vaccine, and the meetings mark macrophages "epigenetically": the configuration of their DNA is changed and passed to daughter cells. These changes enhance immunological responses to several pathogens, not just one, and alter macrophages' metabolism to make them more active defenders. Should a different pathogen attack, the cells produce extra cytokines that trigger inflammation and other bodily processes that harm invaders.


Pathogen A (or vaccine with live weakened pathogen)
Pathogen is ingested and destroyed
Naive macrophage
Cytokines
Epigenetic modification of immunological and metabolic pathways
Pathogen B
High cytokine production from the trained macrophage enhances inflammation and activates other immune cells
Trained macrophage

SOURCES: "TRAINED IMMUNITY: A PROGRAM OF INNATE IMMUNE MEMORY IN HEALTH AND DISEASE," BY MIHAI G. NETEA ET AL., IN *SCIENCE*, VOL. 352; APRIL 22, 2016, AND "INNATE IMMUNE MEMORY: A PARADIGM SHIFT IN UNDERSTANDING HOST DEFENSE," BY MIHAI G. NETEA ET AL., IN *NATURE IMMUNOLOGY*, VOL. 16; JULY 2015

that tag pathogens so that they are easier for phagocytes to find.

Given that BCG was increasing protection to multiple pathogens, it made sense to Netea that the innate immune system might be involved. But conventional thinking held that the innate immune system could not "remember" past immunological encounters, such as stimulation from previous vaccines. The thinking has long been that innate immune cells attack whatever they see and then forget about the battle afterward, like a soldier with amnesia. But these assumptions have been woefully incorrect.

In a paper published in 2012 in the *Proceedings of the National Academy of Sciences USA,* Netea's team found that human immune cells primed by BCG produce four times as much of a key cytokine called IFN-gamma (IFN-γ) and twice as much of the cytokines TNF and interleukin-1 beta (IL-1ß) when later exposed to other pathogens. The cells can initiate these enhanced responses for as long as three months after vaccination, which suggests that the innate immune system can, in fact, remember what it learns. More recently, in 2018, the researchers reported that BCG reprograms human immune cells in ways that help them stave off the yellow fever virus.

Illustration by Jen Christiansen

© 2019 Scientific American

NEWBORN CHILD GETS VACCINATED against tuberculosis in a Guinea-Bissau hospital. Some studies indicate the shot protects against many diseases.

Netea "has really pioneered a new field within innate immunology," says Helen Goodridge, an immunologist at Cedars-Sinai Medical Center in Los Angeles. Studies by other labs also support his theory, showing that the measles vaccine boosts the body's immune response to the toxin produced by tetanus bacteria, as well as its response when exposed to *Candida*.

It is unclear how the measles vaccine elicits its broad effects, but Netea's work suggests that BCG trains the innate immune system by initiating changes in cellular metabolism and by shaping how key immune genes are controlled. After a person gets BCG, little molecular stamps are placed on important immune-related genes, and these stamps later identify the genes so that they can be quickly turned on when another pathogen invades. Why would a live vaccine elicit these effects better than an inactivated one? Researchers theorize that live organisms may stimulate a different reaction simply because they are alive—not just bits and pieces of an organism, as in the inactivated shots. (Real full-on infections, such as measles, do not seem to produce these advantageous effects and can actually suppress immunity.)

While wrapping up his 2012 study in *PNAS,* Netea stumbled across a trial that had just been published by Aaby and Benn suggesting that BCG reduces general neonatal mortality—a finding that was criticized for being biologically impossible. Excited, Netea wrote to Aaby, telling him that he had just discovered a mechanism that made sense of his findings. Since then, the two researchers and Benn have been working together to tease out the immunology behind the Guinea-Bissau data. Vaccines seem to "change the immune system, and they don't just change it in the adaptive, pathogen-specific way," says Tobias Kollmann, an immunologist and infectious disease physician at the University of British Columbia, who sometimes collaborates with Aaby, Benn and Netea. "They change it in all kinds of different ways."

TRIALS ON TRIAL

NEAL HALSEY AGREES that Aaby has made important contributions to vaccine research over the course of his career—but his work on off-target effects is not one of them. Halsey, former director of the Johns Hopkins University's Institute for Vaccine Safety, goes back a long way with the Danish scientist. He remembers that in the 1980s, Aaby was the first to identify a potential safety problem with a new, more concentrated measles vaccine introduced in Guinea-Bissau and other developing countries. At first, no one believed him—this appears to be a recurrent Aaby pattern—but then Halsey looked at data he had collected in Haiti and saw the same effects. Based largely on their findings, the WHO withdrew the vaccine from use in 1992.

But today Halsey thinks that Aaby is putting his convictions before the science. At the 2018 World Vaccine Congress in Washington, D.C., Halsey said the data from Guinea-Bissau may be real, but Aaby and Benn have been drawing causal conclusions from it that they shouldn't. Kids who get vaccinated on time are often quite different from those who don't: they can be healthier to begin with, or they can have wealthier parents with the means to drive them to the doctor and take better care of them in general. Concluding that vaccines are responsible for broadly different medical outcomes is too much of a stretch, Halsey says.

A 2017 *BMJ* study from the Netherlands illustrates his point. Researchers analyzed hospitalization rates among toddlers who had received a live vaccine as one of their last shots and then compared them with hospitalization rates among toddlers who had most recently gotten only inactivated vaccines. Scientists found that the live-vaccinated kids were 38 percent less likely than the others to be hospitalized for infections—but those children were also 16 percent less likely to be hospitalized for injuries or poisoning. Vaccines should not affect accident risk; the fact that the researchers found this link underscores the notion that vaccine history aligns with other factors in one's life. The authors concede, though, that the way vaccines are administered in the Netherlands—they are scheduled in advance, and parents usually cancel appointments only if their kids are sick—most likely inflates the "healthy vaccinee" effect, as it is called, and findings from other countries may not be skewed so heavily.

Because it is so difficult to interpret causality from observational studies, Halsey and others have called for Aaby and Benn to conduct more randomized controlled trials, the so-called gold standard for teasing out an intervention's effects. In these studies, children are randomly selected to receive vaccines or placebos and then followed over time. This random allocation eliminates the chance that socioeconomic status or overall health will play a role in vaccine decisions. The problem is that vaccines are already recommended public policy around the world, so it is unethical

MELINDA WENNER MOYER

© 2019 Scientific American

for researchers to deny children vaccinations to study them. Thus, scientists must get creative—they either have to design trials that provide children with extra vaccines or early ones, or they have to take advantage of natural delays in vaccine receipt.

To undertake a clinical trial in Guinea-Bissau is especially difficult. Aaby and Benn must store vaccines in a refrigerator at their house, where they have a generator, because the electrical grid is so unpredictable (they lost power every day during my visit). Political instability is another problem: one of their attempted trials was disrupted by a devastating civil war in 1998, in which Aaby also suffered a near-fatal wound when he was lanced by a piece of iron left behind by a thief who had looted his house. Some Bissau residents speak only rare dialects, which makes things difficult as well, and many don't have phones.

Despite these challenges, Aaby and Benn are trying randomized trials, such as the one involving Maria. In a few completed tests, the results have not always supported their earlier findings. In a 2018 trial that Aaby and Benn worked on, for instance, researchers found that babies who got the recommended measles vaccine at nine months, plus an additional measles shot between four and 4.5 months, were no less likely to be hospitalized or to die than babies who did not get the extra doses. Yet the two are convinced the vaccine effects are real, just not fully understood. Halsey, though, finds their dogged persistence concerning. "Very good objective scientists acknowledge when an initial observation they made is shown not to be true," he says.

Aaby and Benn are unpopular for another reason: they have published studies suggesting that inactivated vaccines, such as DTP, have detrimental effects, particularly for girls. Even though these vaccines protect against their targeted diseases, Aaby and Benn have linked these shots to a higher risk of other infectious diseases. It is unclear why this would happen—perhaps exposure to dead pathogens makes the immune system more tolerant of other future intruders—and critics argue the associations are not just spurious but also dangerous because they could further undermine the public's confidence in vaccines. "Some of them just think that I'm a madman making trouble," Aaby concedes.

A SEARCH FOR CLARITY

HIS BATTLES, HOWEVER, are entering a new phase. Although Aaby notes that his own research funds are running short, the WHO says that it will soon step into the arena. Aaby first contacted the agency about his findings in 1997; in 2013 it established a working group to review the data. In 2014 the WHO noted that the issue deserved further attention, and in 2016 and 2017 it discussed plans to oversee additional trials. One trial will investigate the effects on infant mortality of giving BCG vaccination at birth versus a placebo. The other will evaluate the effects of an extra dose of measles vaccine given with DTP between 12 and 16 months of age.

Aaby and others worry, however, that these trials will yield little clarity. The subjects will be given inactivated vaccines either at the same time as the live vaccines or after them, which, according to Aaby's previous findings, could mute potentially beneficial effects. "We discussed this at length with many experts, and the evidence is clear that those trials will not give the answer," Kollmann says. Shann, the Australian pediatrician, agrees. These trials will be "a scandalous waste of time and money," he says, because "none of those involved really understands the field." And right now it is unclear when the trials will start. WHO spokesperson Tarik Jasarevic says that as of early 2019, the agency had not found financial sponsors for the work.

Ultimately Aaby worries that the WHO is just going through the motions. He suspects the agency wants to appear that it is doing due diligence after its 2014 report on nontargeted effects but that its real goal is to make the issue go away. If nonspecific effects are real and powerful enough to save lives, then public health agencies will have to consider making changes to the vaccine schedule and perhaps even replace some inactivated vaccines with live ones, which would be extremely difficult.

Last year I asked Frank DeStefano, director of the CDC's Immunization Safety Office, what it would take to make such changes in the U.S. "Certainly evidence would have to be stronger that this is a real effect," he said. He noted that the agency had no plans at that time to collect more data on the issue. But even if it had additional evidence, he said, the CDC would have to consider all the possible risks and benefits before making policy changes.

The evening I left Guinea-Bissau I sat in the back garden with Benn, eating Danish cheese that she brought with her from her last trip home, and I thought about the couple's philosophy of science. These researchers are not shy about their beliefs; they are convinced that nonspecific effects are real but so complex that many details remain a mystery, and they are not afraid to say so. To critics, this strength of conviction is a great weakness, a blazing preconception that biases their results. And it may do so. But bias is not unique to them. Scientists are people—people with ideas, and prejudices, and feelings—and every study involves interpretation. How do we know whose interpretations edge closest to the truth? Are those who admit to their beliefs more biased than those who don't? Who should decide when enough evidence has amassed to reach a consensus, particularly when the implications are unexpected, inconvenient and consequential? Within this small and contentious field, at least, there are no clear answers.

"You have this feeling you are pulling a thread, and you don't know how big the ball of yarn is," Benn said to me. She was referring to the research on vaccines, but she could have been speaking about the scientific process itself. Biology is immensely complicated because our bodies are complex. The practice of science is complicated, too, because it is a product of humanity—an endeavor created and shaped by our imperfect minds. If vaccines do what Aaby and Benn think they do—and that is still an open question—it will take a lot more messy unraveling before the world sees things their way. SA

Contributing editor **Melinda Wenner Moyer** wrote "American Epidemic" (May 2018), which won second place for health policy in the Awards for Excellence in Health Care Journalism.

MORE TO EXPLORE

Vaccine Programmes Must Consider Their Effect on General Resistance. Peter Aaby et al. in *BMJ*, Vol. 344, Article No. e3769. Published online June 14, 2012.

The Introduction of Diphtheria-Tetanus-Pertussis and Oral Polio Vaccine among Young Infants in an Urban African Community: A Natural Experiment. Søren Wengel Mogensen et al. in *EBiomedicine*, Vol. 17, pages 192–198; March 2017.

Trained Immunity: An Ancient Way of Remembering. Mihai G. Netea and Jos W. M. van der Meer in *Cell Host & Microbe*, Vol. 21, No. 3, pages 297–300; March 8, 2017.

BCG Vaccination Protects against Experimental Viral Infection in Humans through the Induction of Cytokines Associated with Trained Immunity. Rob J. W. Arts et al. in *Cell Host & Microbe*, Vol. 23, No. 1, pages 89–100; January 10, 2018.

scientificamerican.com/magazine/sa

© 2019 Scientific American

IS DEATH REVERSIBLE?

An experiment that partially revived slaughterhouse pig brains raises questions about the precise end point of life

By Christof Koch

*"And death shall have no dominion"—**Dylan Thomas, 1933***

You will die, sooner or later. We all will. For everything that has a beginning has an end, an ineluctable consequence of the second law of thermodynamics.

Few of us like to think about this troubling fact. But once birthed, the thought of oblivion can't be completely erased. It lurks in the unconscious shadows, ready to burst forth. In my case, it was only as a mature man that I became fully mortal. I had wasted an entire evening playing an addictive, first-person shooter video game—running through subterranean halls, flooded corridors, nightmarishly turning tunnels, and empty plazas under a foreign sun, firing my weapons at hordes of aliens relentlessly pursuing me. I went to bed, easily falling asleep but awoke abruptly a few hours later. Abstract knowledge had turned to felt reality—I was going to die! Not right there and then but eventually.

Evolution equipped our species with powerful defense mechanisms to deal with this foreknowledge—in particular, psychological suppression and religion. The former prevents us from consciously acknowledging or dwelling on such uncomfortable truths while the latter reassures us

IN BRIEF

Death has had a changing definition over the millennia. Originally, it meant cessation of breathing and a heart that had stopped.

The advent of mechanical ventilators shifted the locus of death to the brain—dying became loss of brain function, an irreversible coma.

Partial revival of pig brains hours after decapitation, which was demonstrated in a recent experiment, could again upend definitions of mortality.

Illustration by Bill Mayer

© 2019 Scientific American

© 2019 Scientific American

by promising never-ending life in a Christian heaven, an eternal cycle of Buddhist reincarnations or an uploading of our mind to the Cloud, the 21st-century equivalent of rapture for nerds.

Death has no such dominion over nonhuman animals. Although they can grieve for dead offspring and companions, there is no credible evidence that apes, dogs, crows and bees have minds sufficiently self-aware to be troubled by the insight that one day they will be no more. Thus, these defense mechanisms must have arisen in recent hominin evolution, in less than 10 million years.

Teachings from religious and philosophical traditions have long emphasized the opposite: look squarely into the hollow eyes of death to remove its sting. Daily meditation on nonbeing lessens its terror. As a scientist with intimations of my own mortality, my reflections turn toward understanding what death is.

Anyone who undertakes this quest will soon come to realize that death, this looming presence just over the horizon, is quite ill defined from both a scientific as well as a medical point of view.

FROM THE CHEST TO THE HEAD

THROUGHOUT HISTORY, everyone knew what death was. When somebody stopped breathing and his or her heart ceased beating for more than a few minutes, the person was, quite simply, dead. Death was a well-demarcated moment in time. All of this changed with the advent of mechanical ventilators and cardiac pacemakers in the middle of the 20th century. Modern high-tech intensive care decoupled the heart and the lungs from the brain that is responsible for mind, thought and action.

In response to these technological developments, in 1968, the famous *Report of the Ad Hoc Committee of the Harvard Medical School* introduced the concept of death as irreversible coma—that is, loss of brain function. This adjustment was given the force of law by the Uniform Determination of Death Act in 1981. This document defines death as either irreversible cessation of circulatory and respiratory functions or irreversible halting of brain function. Quite simply, when your brain is dead, you are dead.

This definition is, by and large, in use throughout most of the advanced world. The locus of death shifted from the chest to the brain (and from public view into the private sphere of the hospital room), with the exact time of actual brain death uncertain. This rapid and widespread acceptance of brain death, reaffirmed by a presidential commission in 2008, is remarkable when compared with the ongoing controversy around abortion and the beginning of life. It may perhaps be reflective of another little noticed asymmetry—people agonize about what happens in the hereafter but rarely about where they were before being born!

The vast majority of deaths still occur following cardiopulmonary cessation, which then terminates brain functioning as well. Neurological death—specified by irreversible coma, absence of responses, brain stem reflexes or respiration—is uncommon beyond the intensive care unit, where patients with traumatic or anoxic brain injury or in a toxic-metabolic coma (say, following an opioid overdose) are typically admitted.

Brain death may be the defining factor, but that does not simplify clinical diagnosis—biological processes can persist after the brain shuts down. Indeed, a brain-dead body can be kept "alive" or on "life support" for hours, days or longer. For the grieving relatives and friends, it is challenging to understand what is happening. When visiting the ICU, they see the chest moving in and out, they feel a pulse, the skin pallor looks normal, and the body is warm. Looking healthier than some of the other denizens of the ICU, their beloved is now legally a corpse, a beating-heart cadaver. The body is ventilated and kept suspended in this quasi-living state because it is now a potential organ donor. If permission has been obtained, the organs can be harvested from the cadaver to help the living who need a heart, kidney, liver or lung, which are always in short supply.

Brain-dead bodies can continue to grow fingernails, to menstruate, with at least some working immune function that allows them to fight off infections. There are more than 30 known cases of pregnant brain-dead mothers placed on a ventilator to support gestation of a surviving fetus, born weeks or months (in one case 107 days) after the mother became brain-dead. In a widely discussed 2018 story in the *New Yorker,* a young woman, Jahi McMath, was maintained on ventilation in a home care setting in New Jersey by her family following her brain death after tonsil surgery in a California hospital. To the law and established medical consensus, she was dead. To her loving family, she was alive for nearly five years until she died from bleeding associated with liver failure.

Despite technological advances, biology and medicine still lack a coherent and principled understanding of what precisely defines birth and death—the two bookends that delimit life. Aristotle wrote in *De anima* more than two millennia ago that any living body is more than the sum of its parts. He taught that the vegetative soul of any organism, whether a plant, animal or person, is the form or the essence of this living thing.

The essence of a vegetative soul encompasses its powers of nutrition, growth and reproduction that depend on the body. When these vital capacities are gone, the organism ceases to be animate (a term whose roots lead back to *anima,* Latin for "soul"). The sensitive soul mediates the capacities of both animals and humans to sense the world and their bodies. It is the closest to what we moderns call "conscious experience." Finally, the rational soul is the sole province of people, mediating reason, language and speech. Of course, this is now increasingly mimicked by artificial-intelligence algorithms.

The modern emphasis on machine learning, genomics, proteomics and big data provides the illusion of understanding what this vegetative soul is. Yet it obscures the depth of our ignorance about what explains the breakdown of that entity. A conceptual challenge remains to define what constitutes anyone's living body—which is clearly more than the sum of its individual organs. How can one precisely delimit this body in space (are clothing, dental implants and contact lenses part of the body?) and in time (its beginning and its end)?

Note the word "irreversible" in the contemporary definition of neurological death. In the absence of a precise conceptual formulation of when an organism is alive or dead, the concept of irreversibility depends on the technology du jour, which is constantly evolving. What at the beginning of the 20th century was irreversible—cessation of breathing—became reversible by the end of the century. Is it too difficult to contemplate that the same may be true for brain death? A recent experiment suggests this idea is not just a wild imagining.

PARTIAL REVIVAL OF DEAD BRAINS

THIS YEAR a large team of physicians and scientists at the Yale School of Medicine under Nenad Sestan took advantage of hun-

© 2019 Scientific American

dreds of pigs killed at a Department of Agriculture–approved slaughterhouse for a remarkable experiment, published in the journal *Nature*. The researchers removed the brains from their skulls and connected the carotid arteries and veins to a perfusion device that mimics a beating heart. It circulates a kind of artificial blood, a synthetic mixture of compounds that carry oxygen and drugs that protect cells from damage. The magic resides in the exact molecular constitution of the circulating solution. Think of closed-circuit dialysis machines that thousands of patients use daily to flush out toxins from their body because their own kidneys have stopped working.

These machines are needed because when blood stops flowing through the large, energy-demanding brain, oxygen stores are depleted within seconds, and consciousness is lost. Depriving a brain of oxygen and blood flow for more than a few minutes begins to trigger irreversible damage. Cells start degenerating in all sorts of ways (tissue damage and decomposition, edema, and so on) that are readily visible under a microscope.

The Sestan team studied the brains' viability four hours after the pigs were electrically stunned, bled out and decapitated. (If this sounds gruesome, it is what happens to livestock in an abattoir, one reason I'm a vegetarian.) The researchers compared a variety of biological indicators with those of postmortem control brains from pigs that did not undergo this perfusion procedure four hours after death, an eternity for the sensitive nervous system.

At first glance, the restored brains with the circulating solution appeared relatively normal. As the compound circulated, the fine net of arteries, capillaries and veins that suffuse brain tissue responded appropriately; the tissue integrity was preserved with a reduction in swelling that leads to cell death; synapses, neurons and their output wires (axons) looked normal. Glial cells, the underappreciated entities supporting neurons proper, showed some functionality, and the brain consumed oxygen and glucose, the universal energy currency of the body, an indication of some metabolic functioning. The title of the researchers' paper announcing their technology boldly states "Restoration of Brain Circulation and Cellular Functions Hours Post-mortem."

What was not present in these results were brain waves of the kind familiar from electroencephalographic (EEG) recordings. Electrodes placed onto the surface of the pig brains measured no spontaneous global electrical activity: none of the slow waves that march in lockstep across the cerebral cortex during deep sleep, no abrupt paroxysm of electrical activity followed by silence—what is known as burst suppression. Only a flat line everywhere—a global isoelectric line—implying a complete absence of any sort of consciousness. A silent brain, electrically speaking, is not harboring an experiencing mind. But this was not a surprise. This state was exactly what was intended by Sestan and his co-workers, which is why the circulating solution contained a cocktail of drugs that suppresses neuronal function and corresponding synaptic communication among cells.

Even with the absence of brain waves it came as a surprise to me, a working neuroscientist, that individual pig cortical neurons still retained their capacity to generate electrical and synaptic activity. The Yale team demonstrated this by snipping a tiny sliver of neural tissue from these brains, washing off the perfused solution and then exciting individual neurons via an electric current delivered by a tiny electrode. Some of these cells responded appropriately by generating one or a series of the stereotypical electrical pulses, so-called action potentials or spikes, that are the universal idiom of rapid communication in any advanced nervous system.

This finding raises a profound question: What would happen if the team were to remove the neural-activity blockers from the solution suffusing the brain? Most likely nothing. Just because some individual neurons retain some potential for excitability does not imply that millions and millions of neurons can spontaneously self-organize and break out into an electrical chorus. And yet! It cannot be ruled out that with some kind of external help, a sort of cortical defibrillator, these "dead" brains could be booted up, reviving the brain rhythms characteristic of the living brain.

To state the obvious, decapitating any sentient creature and letting its brain bleed out is not conducive to its well-being. Reanimating it after such a major trauma could well lead to profound pathology, such as massive epileptic seizures, delirium, deep-seated pain, distress, psychosis, and so on. No creature should ever suffer in this manner. It is precisely to avoid this situation that the Yale team obstructed neuronal function.

This brings me to the elephant in the room. Can this procedure be applied to the human brain? Before you recoil, think of the following. What would you want done if your child or partner were found drowned or overdosed, without a pulse or breath for hours? Today it is likely that they would be declared dead. Could this change tomorrow with the kind of technology pioneered by the Yale group? Isn't that a worthwhile goal to pursue?

The pig brain is a large brain, unlike the one of the much smaller mouse, by far the most popular laboratory animal. Pig cortex is highly folded, like the human cortex. Neurosurgical procedures are routinely tested on pigs before moving to human trials. So, the technical answer is yes; in principle, this could be done.

But should it be done?

Certainly not until we understand much better whether a reconstituted animal brain shows global electrical activity typical of a healthy brain, without stress responses indicative of pain, distress or agony. The field as a whole should pause and discuss the medical, scientific, legal, ethical, philosophical and political questions of such research with all stakeholders.

Yet the fear of the grim reaper will not be denied. Sooner or later, somewhere on the planet's face, someone will try to temporarily cheat death. SA

Christof Koch is chief scientist and president of the Allen Institute for Brain Science in Seattle. He serves on *Scientific American*'s board of advisers and is author of *The Feeling of Life Itself: Why Consciousness Is Widespread but Can't Be Computed* (MIT Press, 2019).

MORE TO EXPLORE

The Undead: Organ Harvesting, the Ice-Water Test, Beating-Heart Cadavers—How Medicine Is Blurring the Line between Life and Death. Dick Teresi. Pantheon, 2012.

Modern Death—How Medicine Changed the End of Life. Haider Warraich. St. Martin's Press, 2017.

Restoration of Brain Circulation and Cellular Functions Hours Post-mortem. Zvonimir Vrselja et al. in *Nature,* Vol. 568, pages 336–343; April 18, 2019.

scientificamerican.com/magazine/sa

© 2019 Scientific American

Distant New World

Ultima Thule, the farthest object ever visited by a spacecraft, is revealing our solar system's deepest history—and, just maybe, revolutionizing planetary science

By Corey S. Powell

© 2019 Scientific American

THIS IS THE MOST detailed image of Ultima Thule returned so far by the New Horizons spacecraft. It was taken in the early hours of January 1, 2019, just 30 minutes before its closest approach from a range of 18,000 miles (28,000 kilometers). Further studies of this image and others are set to lift the veil on hidden chapters of our solar system's history.

NASA/JOHNS HOPKINS APPLIED PHYSICS LABORATORY/SOUTHWEST RESEARCH INSTITUTE, NATIONAL OPTICAL ASTRONOMY OBSERVATORY

© 2019 Scientific American

The history of the solar system is a dish best served cold. And it is so very cold on Ultima Thule.

That is the message that was beamed back to Earth from NASA's New Horizons probe when it completed its historic exploration of a small body in the Kuiper belt, the sprawling population of dwarf planets and cometlike objects out beyond the orbit of Neptune. When New Horizons flew past at 12:33 A.M. Eastern Time on January 1, 2019, Ultima was a hair over four billion miles from the sun. It is by far the most distant object ever visited by spacecraft and correspondingly one of the coldest: about 35 kelvins, or nearly 400 degrees below zero Fahrenheit.

At such low temperatures, Ultima (more formally known by its scientific designation, 2014 MU69) preserves its initial, ancient composition. Ultima is also cold in another, more specialized and intriguing way. It is *dynamically* cold, part of what's known as the "cold classical" Kuiper belt, meaning that it circles the sun in a settled orbit that was undisturbed by all the chaotic events that buffeted Earth and other planets as they came together more than four billion years ago.

Prior to the encounter, the great hope of the New Horizons team was that the scientists would see an intact survivor from the solar system's birth. The first images of Ultima, showing its delicately stacked snowman shape, fully vindicate those hopes. It matches up exactly with models of how clouds of gas and dust around young stars clump together into larger and larger objects—a process that has been well studied in theory but never observed in reality until now. "We're looking at one of the first building blocks that came together to form the planets and moons," says Jeffrey Moore, a research scientist at NASA's Ames Research Center. "It looks like somebody left it out in the back of God's freezer for the past four and a half billion years."

LONG WAY TO THE KUIPER BELT

BEING LEFT OUT IN THE COLD is an all too familiar feeling for Alan Stern, principal investigator of New Horizons. He had started campaigning for NASA to mount a mission like the one to Ultima in the 1980s, long before anyone even knew the object existed. His original goal was to visit Pluto, completing the exploration of the then nine planets, but the plans went nowhere. Befitting their outsider status, Stern and other like-minded fans of the outer solar system called themselves the "Pluto Underground."

Scientific interest in Pluto exploded in the 1990s, when astronomers discovered it is just one part of the much grander Kuiper belt. Even so, Stern watched three different mission concepts wither and die. It took another decade, a cancellation and a rare resurrection before New Horizons launched on January 19, 2006.

The probe's nominal objective was only to visit Pluto and its moons, but from the start Stern had a second act in mind. "New Horizons is a very healthy spacecraft, and it has the power to operate for another 15 or 20 years," he says. Why waste a once-in-a-lifetime opportunity to explore more of the Kuiper belt? A 2015 flyby revealed Pluto as an excitingly dynamic world, but in some ways it is too dynamic: its active geology (Plutology?) has erased much of its early history. A visit to a much smaller, more pristine follow-on target would fill in a huge missing piece of the reconnaissance of the solar system.

The catch was that there was no suitable target in astronomers' catalogs; the New Horizons team would have to find one itself, something the spacecraft could reach on its dwindling stores of fuel. That meant marshaling the world's most powerful telescopes, searching for a dim dot of light moving slowly among a thicket of background stars, located in exactly the right spot for an easy interception. Marc Buie of Southwest Research Institute (SwRI) describes his work leading that search. "I spent 10 years looking for an object for New Horizons to go to," he says with admirable sangfroid. "I'll let you in on a little secret: when you take photographs through a telescope, they don't come with arrows on them to tell you what to look at."

Finally, Buie spotted it. The telltale blip consisted of just 144 photons in a survey image from the Hubble Space Telescope, "but the instant I saw it, I got a chill. I knew it was the One." As recognition of its importance increased, the object originally catalogued as 1110113Y became Potential Target 1, then Kuiper belt object 2014 MU69, and most recently it was informally dubbed Ultima Thule.

ENTER SNOWMAN

WHEN THE FIRST HIGH-RESOLUTION images of Ultima arrived from New Horizons, all those past agonies vanished from the scientists' minds, replaced by backslaps and oversize, unshakable grins.

"It's a snowman!" Stern laughed as he looked at Ultima's iconic, double-lobed shape. Or in scientific terms, it is a "contact binary," two objects that formed separately but are now held together by their mutual gravity. The larger lobe is three times the size of the smaller one; stacked together, they measure 21 miles long.

That snowman structure is precisely what the New Horizons team was hoping to find in the Kuiper belt. "The number-one thing I wanted to get out of this mission was incontrovertible evidence that we're looking at a primordial, unaltered object," Buie says. More specifically, he wanted to witness an object frozen in time from the moment when small bodies in the solar system began accreting together and building up into bigger ones.

On planets and moons all evidence of that process was wiped clean by subsequent geologic activity. Even on relatively primitive comets such as 67/P Churyumov-Gerisamenko—the "rubber duck" comet studied until 2016 by the European Space Agency's Rosetta probe—the surface has been extensively cooked and altered by the sun's heat. Moore breezily dismisses such objects as "post-toasties."

On Ultima it is as if the birth of the solar system happened yesterday, with two planetesimals (the first large conglomerations of rock, dust and ice to coalesce around a young star) still lumped together right before New Horizons's digital eyes. "I looked at it and thought, 'I see accretion happening,'" Buie says. "This is going to revolutionize our view of where we came from and how this whole process works."

© 2019 Scientific American

DO YOU WANT TO BUILD A PLANET?

IN FACT, THE REVOLUTION is already well underway. On the first three days after the Ultima flyby, the members of the New Horizons analysis teams huddled near mission control at Johns Hopkins University Applied Physics Laboratory to make sense of the first data downloads. After all those decades of waiting, Stern is eager to get going. "We will begin writing our first science papers next week!" he promises. But really, nobody was waiting even that long to start digging into the lessons from Ultima.

One of the first orders of business is interpreting how Ultima's two lobes came together so gently, with a tidy little necklace of brighter material rimming the place where they meet. Bill McKinnon, a planetary dynamics expert at Washington University in St. Louis, notes that objects in the Kuiper belt are moving so slowly that typical encounter velocities would be around 600 miles per hour, which limits the amount of damage an impact can create.

Ultima apparently came together even more lightly than that. Most likely the two lobes initially formed as separate bodies in orbit around each other, spiraling together until they touched. "It would have been like a parking lot fender bender, walking speed. Also, these are probably very porous objects; they have energy-absorbing bumpers on them," McKinnon says. As a result, the two lobes stuck together rather than shattering when they touched.

The jammed-together lobes of Ultima provide the most direct evidence yet of the standard hierarchical theory of how planets form: First, dust grains stick together into pebbles, which gather into planetesimals. Those accrete into protoplanets, continuing all the way up to bigger things like Earth and Jupiter.

Another notable detail is that, other than the difference in size, the two lobes of Ultima look extremely similar in shape and color. "That's also consistent with forming as a result of the merger of two objects that formed close together in the protoplanetary disk," says Silvia Protopapa of SwRI. In such a chemically and dynamically cold region, two neighboring planetesimals should have aggregated from an essentially identical mix of raw materials.

Information still stored onboard New Horizons will reveal much about exactly what those materials are. "We'll be looking for ammonia, water, carbon monoxide and organics," Protopapa says. Ultima's overall red tint already suggests it is covered with tholins (organic compounds processed by eons of radiation), similar to what is seen on many other outer solar system objects, including Pluto's moon Charon.

Moore is relieved, if maybe a little disappointed, that nothing about Ultima directly confounds expectations—as would be the case if, say, the two lobes had totally different colors. "To zeroth order, our ideas of how these things formed seem to be vindicated," he says. "But there's a whole range of specific explanations that people had. It's like going into a restaurant—you can choose from any number of things on the menu. Now that we see Ultima, we know what the good choices on the menu are."

ULTIMA'S NEXT ACT

THESE RAPID-FIRE READINGS of Ultima are all the more impressive considering they were made by highly caffeinated scientists operating on little or no sleep, working with an initial download of just 1 percent of the total data New Horizons collected during its flyby.

Perhaps the most egregious missing data element so far is a clear look at the topography of Ultima: we can see it, but we can't *really* feel it. The best views released so far were taken with the sun almost directly behind the spacecraft, making it difficult to distinguish hills from depressions on the alien landscape below. Significantly sharper pictures, including ones showing deep shadows, are still sitting on New Horizons's solid-state recorders. "It'll take 20 months to empty the recorders of all the data we've taken, including hundreds of images and spectra," Stern says. Blame the extremely limited bandwidth of a spacecraft six light-hours from home.

As the stored information trickles in, you can expect a lot more history lessons. One big hanging question for the planetary scientists is how Ultima's individual lobes themselves were assembled. Somewhat surprisingly, there's no unambiguous evidence for craters. Certainly there are things that resemble craters, but in the calm environment where Ultima formed, those poorly resolved formations might actually be hills that were built up rather than gouged-out pits. "We may find that the texture is dominated by slow-motion [accretionary] processes," Moore says. If so, that would reveal a lot about the early stages of planetary formation.

With that thought in mind, Cathy Olkin of SwRI, a deputy project scientist on New Horizons, is eager to examine how the composition of Ultima varies from spot to spot. Spectral measurements from New Horizons should make it possible to identify the individual splats from ancient impacts that stuck to the surface like flies on a windshield. Then she'll do a census of the splats to understand the objects that created them. "So we will be able to look at the smallest particles in the Kuiper belt, a population we could not see any other way," she says.

Perhaps the most subtle but significant surprise about Ultima is that we were able to reach it at all. Recent studies of newborn stars using the Atacama Large Millimeter/submillimeter Array (ALMA) observatory in Chile show that many of them seem to be forming planets across a much larger scale than what we see in our solar system, with their versions of the Kuiper belt taking shape at far greater distances. The vast distance between our sun and Ultima does not look so vast in comparison.

McKinnon throws out another possible history lesson: the outer edge of our solar system may have been clipped off by radiation from neighboring stars at a very early stage, he suggests, well before even Ultima formed. "It's probably defined by the other stars in the sun's birth cluster, its lost brothers and sisters," he says.

In the debate over whether our solar system is normal or a cosmic outlier, then, Ultima tips the balance a tiny bit further in the weirdo direction. "Then you consider that the sun is a single star, which is not the norm, and solar-type stars are not the most common stars," McKinnon says. "You add it up, and you think, 'We're not exactly a garden-variety planetary system.'" SA

Corey S. Powell is a science writer, blogger and editor living in Brooklyn, N.Y. He is a visiting scholar at New York University's Science, Health and Environmental Reporting Program.

MORE TO EXPLORE

20 Years Later—a Q&A with the First Astronomer to Detect a Planet Orbiting Another Sun. Corey S. Powell; ScientificAmerican.com, October 6, 2015.
Have We Mismeasured the Universe? Corey S. Powell; February 6, 2019.
The Hunt Is On for Alpha Centauri's Planets. Corey S. Powell; August 5, 2019.

scientificamerican.com/magazine/sa

© 2019 Scientific American

Finally, the Face of a Black Hole

After more than a decade of effort, a global network of radio telescopes has revealed the first-ever picture of an enigmatic hole in spacetime

By Seth Fletcher

IT ACTUALLY WORKED.

The Event Horizon Telescope (EHT) has always sounded like an unlikely proposition: create a virtual Earth-size telescope to take a picture of the shadow of a black hole. Every piece of that sentence is a little bit insane.

But it worked. After more than a decade of technical development and fundraising and astronomer herding and telescope wrangling, the astronomers of the EHT unveiled the first picture of a black hole—a bubble of pure gravity, a hole in spacetime, a prediction of general relativity so strange that Albert Einstein himself long refused to believe the concept was possible.

In six simultaneous press conferences on four continents, astronomers lit up projection screens with an image that, for experts at least, is unmistakable: a dark circle surrounded by a ring of light and lopsided fiery blobs, clumps of matter heated to 100 million degrees and circling the black hole at close to the speed of light.

Technically, this is a picture of the *shadow* of a black hole: specifically, a supermassive black hole—a 6.5-billion-solar-mass beast—55 million light-years from Earth at the center of the galaxy Messier 87. Black holes trap everything that falls in, including light, which is why they are black. In a sense, they are fundamentally unseeable, one-way cosmic escape hatches leading to … well, that part is unclear. But because of the way they warp spacetime, they impose a dark silhouette on the glowing, superheated matter that circles them. This is the barely imaginable scene the EHT captured during an observing run two years ago and revealed to the world now.

Since the 1960s, when indirect astronomical evidence and breakthroughs in theoretical physics made the existence of black holes all but undeniable, these objects have been in a kind of epistemological limbo: they were the most likely explanation for all sorts of otherwise inexplicable phenomena, but no one held out much hope of ever seeing one.

That changed in the late 1990s, when astronomers around the world began building powerful high-frequency radio observatories on high ground in Hawaii, Mexico, Chile and the South Pole. Theoretical astrophysicists calculated that, in theory, uniting these telescopes using a technique called Very Long Baseline Interferometry would allow them to see the shadows of supermassive black holes that were sufficiently huge and close to Earth. The biggest prize has always been Sagittarius A*, the four-million-solar-mass black hole at the center of the Milky Way. Sagittarius A* remains unseen—at least by the public. But M87, the only other known black hole within the EHT's grasp, has been captured.

Usually big scientific discoveries like this leak early, but this time the astronomers kept surprisingly tight control. France Córdova, director of the National Science Foundation—which has invested $28 million in the EHT over the years—saw the image for the first time at the press conference on the morning of April 10, 2019. It moved her to tears.

The NSF and other funding institutions promised that the EHT's result would be "groundbreaking," and it certainly was—the first direct evidence for the existence of event horizons, the defining feature of a black hole. But anyone hoping that this image would chip away at Einstein's ironclad theory of general relativity came away disappointed: the black hole looks pretty much exactly like the century-old theory says it should.

But this photograph is only a start. Since the 2017 observation, the EHT array has already grown in size and acuity, with a new telescope joining and another on the way and with the installation of ever more powerful digital-signal-processing hardware at each site. As Sera Markoff, an astrophysicist at the University of Amsterdam, said at the press conference, in the future all kinds of new telescopes will join the EHT to observe M87, Sagittarius A* and other black holes simultaneously in many different bands of light—radio, x-ray, gamma ray, infrared and optical.

It is also worth noting that in the two hours after the press conference, at least six scientific papers on the observation appeared online. They almost certainly contain clues and new questions that will take more time to process than a 24/7 news cycle can tolerate. For now, though, it is worth pausing for a moment to consider the strangeness of nature and the remarkable fact that these sentient, tool-using bipeds on a small world in a backwater solar system somehow managed to turn their planet into a telescope and take a picture of an exit chute from the universe. SA

Seth Fletcher is chief features editor at *Scientific American*. His book, *Einstein's Shadow* (HarperCollins, 2018), is about the Event Horizon Telescope and the quest to take the first picture of a black hole.

MORE TO EXPLORE

Dark Star Diaries. Cosmology blog. Seth Fletcher. ScientificAmerican.com

scientificamerican.com/magazine/sa

EVENT HORIZON TELESCOPE (CC BY 4.0 [HTTPS://CREATIVECOMMONS.ORG/LICENSES/BY/4.0])

© 2019 Scientific American

FIRST-EVER IMAGE of the shadow of the supermassive black hole at the heart of the galaxy Messier 87.

© 2019 Scientific American

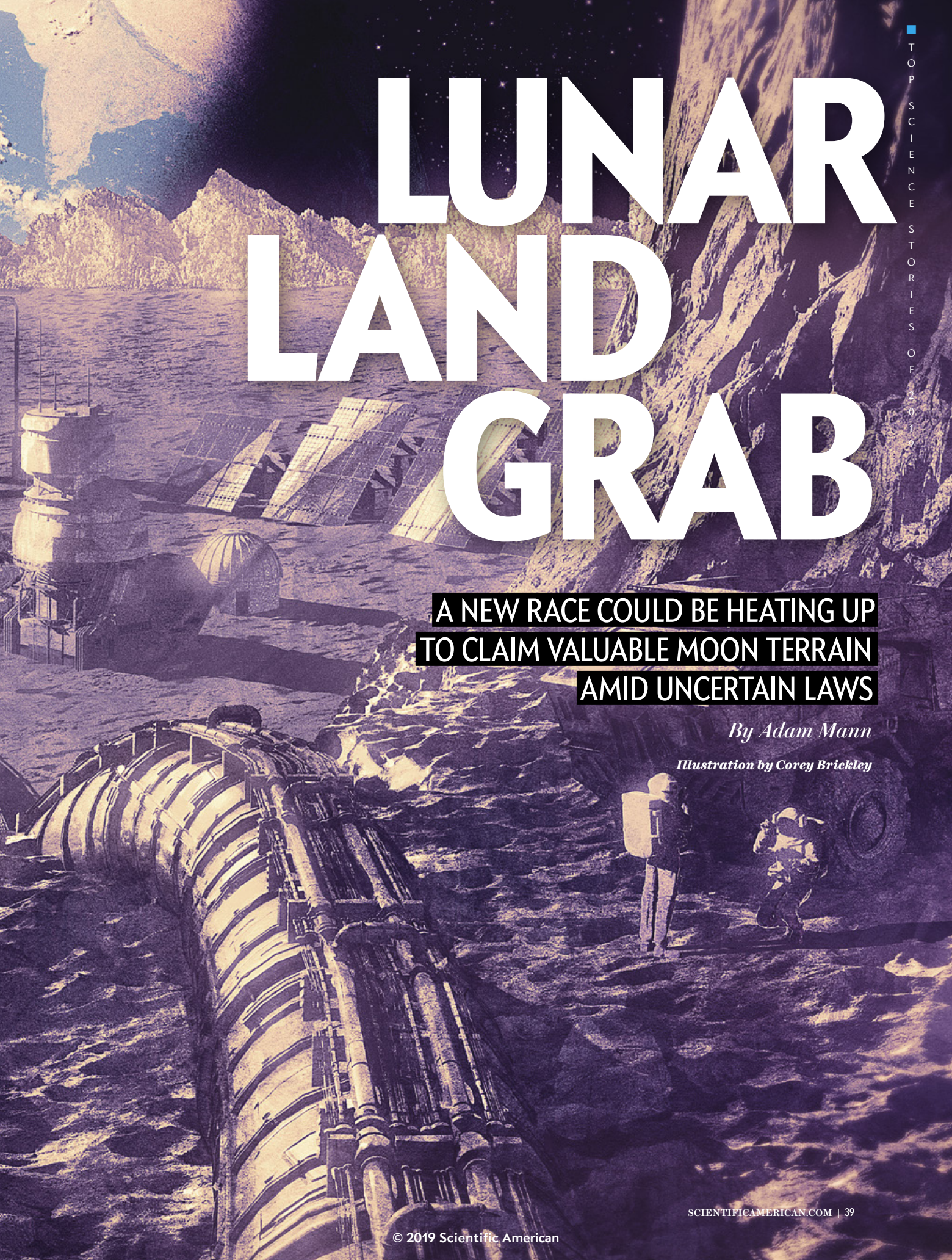
© 2019 Scientific American

LUNAR LAND GRAB

A NEW RACE COULD BE HEATING UP TO CLAIM VALUABLE MOON TERRAIN AMID UNCERTAIN LAWS

By Adam Mann

Illustration by Corey Brickley



© 2019 Scientific American

A GRAINY BLACK-AND-WHITE IMAGE PLAYS ACROSS THE SCREEN in one of Bob Richards's earliest memories—spacesuits, a lander and astronauts Neil Armstrong and Buzz Aldrin taking their historic first steps across the lunar surface. Richards, who was barely out of his toddler years at the time, recalls sitting in his family living room north of Toronto while his dad futzed with the rabbit-ear antenna, trying to improve the broadcast streaming over from Buffalo, N.Y. "*Apollo 11* was a defining moment for humanity," says the founder and CEO of Moon Express, a company that hopes to sell transportation to our natural satellite and eventually mine materials there. "The inspiration of *Apollo* is very prominent in what's happening today in space."

Back in the 1960s, it seemed like just a matter of time before humanity would slip the bonds of Earth and begin a slow crawl out into the universe. Although it has taken longer than many expected, something like that moment may soon arrive. Around half a dozen governments, as well as a handful of private companies, all have moon missions planned for the near future—a situation ripe for conflict.

The Outer Space Treaty, which the U.S., the U.K. and the Soviet Union signed less than two years before *Apollo 11* (and which now has 109 countries party to it), stipulates that space exploration must be conducted peacefully and for the benefit of all nations. It also holds that no one can claim territory on a celestial body. But lower down in the treaty is a loophole: two "noninterference clauses," which require all signatories to avoid causing harm to another's probes or outposts—for instance, by landing near or on top of them. This sounds reasonable enough, but it also creates an opening for a nation or private entity to monopolize a desirable spot simply by arriving there first.

Should one nation or entity try to stake a claim, it "might trigger a 'scramble for the Moon' comparable in some respects to the 'scramble for Africa' which began with the identification of mineral resources in the Congo in the 1880s," wrote astrophysicist Martin Elvis of the Center for Astrophysics at Harvard University and the Smithsonian Institution and his co-authors in a 2016 paper in the journal *Space Policy*.

Sure enough, several missions scheduled to take place in the next few years all target the same territory. India's Chandrayaan-2 mission, launched this past July, was aiming for the lunar poles. The China National Space Administration has said that at least its next three probes will head to the poles as well. The Russian space agency Roscosmos is developing its Luna-Glob program, which would touch down near the Boguslawsky crater near the south pole perhaps as early as 2021. That same year Japan intends to launch the Smart Lander for Investigating Moon, or SLIM, which

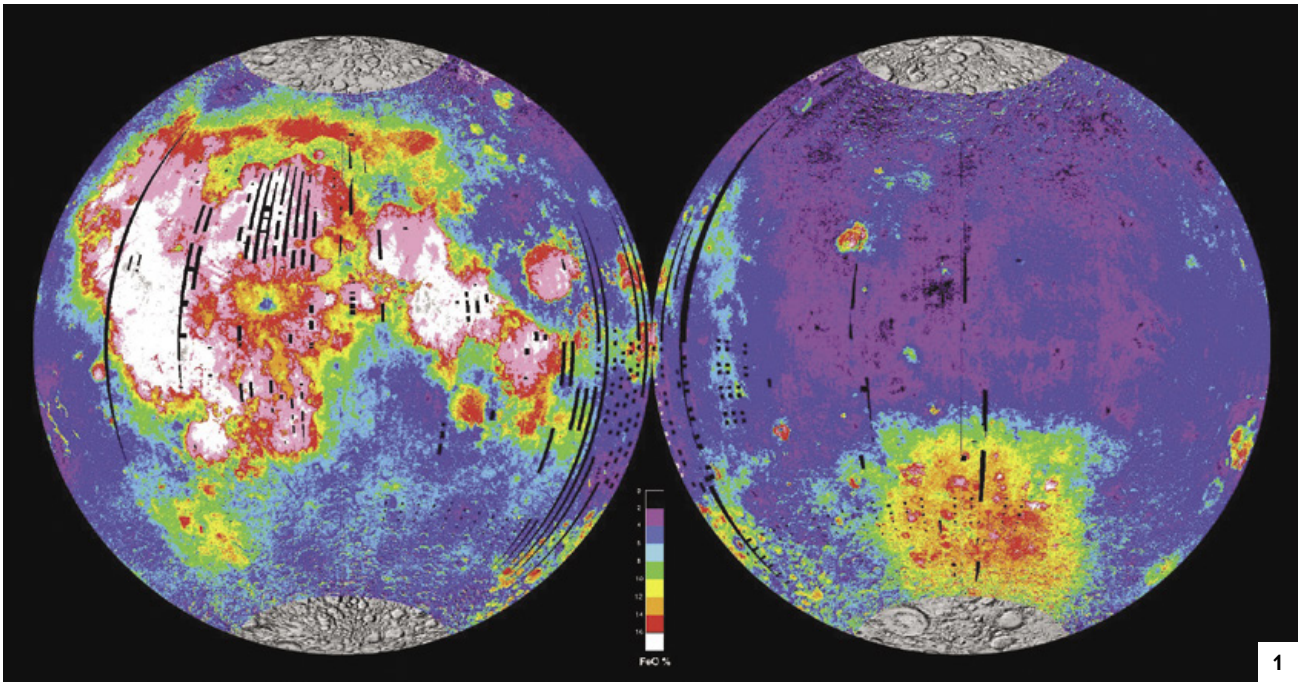
IN BRIEF

A large number of countries and private companies are aiming to launch missions to the moon in the coming decade.

International law says no one can own property in space—yet it also says that once an entity has landed somewhere, others should avoid disturbing that site.

This loophole creates the potential for a race to stake claims on some of the moon's highest-value real estate.

© 2019 Scientific American



1

2

could demonstrate extremely high landing accuracy on small lunar features. NASA, the European Space Agency and more private interests are looking moonward as well. In May, Amazon CEO Jeff Bezos, who founded the spaceflight company Blue Origin, unveiled plans for its *Blue Moon* lunar lander, which, he said, could be ready to carry crews within the next five years.

Moon Express aims to land at the lunar south pole in 2021. And if its spacecraft arrives before anyone else, Richards says, the company wants it to count for something. "One of our drivers is to get there first," he says. "And we expect our rights of noninterference will be respected."

THE OIL OF SPACE

THERE IS PLENTY OF REAL ESTATE on the moon to go around—the total surface area is about the size of Africa—but the resources there are unevenly distributed. Iron and titanium, which could be useful for building moon habitats and technologies, are abundant in different regions of the lunar surface. The helium 3 deposits common in areas of the top layer of lunar regolith could power fusion reactors. And "resources" are not limited to extractable materials. Some landforms, such as certain crater pits, could offer radiation protection to astronauts, and sites on the lunar far side that are shielded from Earth's radio noise would be especially well suited to hosting telescopes.

In the near term, the most desirable resource of all is water. Astronauts can drink water, or they can break it into its constituent elements and transform them into rocket fuel. For the first off-planet explorers, water has been called the oil of space.

Some of the most promising sites for water extraction are the so-called Peaks of Eternal Light at the north and south lunar poles. These are crater peaks, geographical features that often form at or near the edges of impact craters when an asteroid strikes the surface and pushes material to the side, where it rises up to form a ridge at the rim. Because of the moon's orbital mechanics, the sun shines almost perpetually at these peaks, offering a nearly constant source of energy to solar panels. Astronauts could stage bases here to extract the water sitting conveniently nearby at the bottom of these craters, where permanently shadowed regions have allowed ice to accumulate.

Each pole contains roughly half a dozen of these Peaks of Eternal Light, which are about a few hundred meters across apiece. Given this relative scarcity, it is easy to see why the principle of noninterference could be a useful way for nations to claim territory. "They're so small no one else can land on one without risking dam-

IRON concentration on the moon (*1*), as mapped by the Clementine spacecraft in 1994. A mock-up (*2*) of the *Blue Moon* lunar lander being developed by Blue Origin.

NASA (*1*); SAUL LOEB *Getty Images* (*2*)

© 2019 Scientific American

age to a spacecraft that's already there," says Ian Crawford, a planetary scientist at Birkbeck, University of London, who has studied lunar resources. "The first companies or nations that land on these peaks, regardless of the legal niceties, will de facto have ownership."

LEGAL LOOPHOLES

THE OUTER SPACE TREATY was written half a century ago, mainly by two countries—the U.S. and U.S.S.R.—that at the time were the only ones that could even dream of reaching the moon. Legal scholars have debated the treaty's implications ever since, and recent developments, such as the rise of commercial spaceflight, have raised issues that were not on anyone's radar back then.

In 2015 Congress sparked an international disagreement by passing the U.S. Commercial Space Launch Competitiveness Act, which specified that although no one can claim property on a celestial body, any material extracted from one is legally owned by the entity that did the mining and can therefore be sold for profit. Representatives from Russia, Brazil and elsewhere subsequently made an uproar at a March 2017 meeting of the United Nations Committee on the Peaceful Uses of Outer Space (COPUOS), stating that off-world mining was a type of de facto appropriation and that a global consortium was needed to regulate extraction operations.

The Outer Space Treaty was written half a century ago, mainly by two countries—the U.S. and U.S.S.R.—that at the time were the only ones that could even dream of reaching the moon.

Other countries have sided with the U.S.—including nations such as Luxembourg, which would like to be a major player in space-based resources—in its more laissez-faire interpretation of the law. Chinese delegates at COPUOS have mostly threaded the needle between the two sides, appearing to wait and see which reading will eventually prevail. "International law is made by the states collectively," says Frans von der Dunk, a law professor specializing in space at the University of Nebraska–Lincoln. "If one says this is legal and another says this requires an international regime and licensing, then you have a big problem on your hands."

As things stand today, highly desirable areas of the moon are likely to be acquired on a first-come-first-served basis, rewarding the wealthy countries and companies that can get there soon. Less affluent nations could store up animosity toward those that stake claims, stoking tensions much like in the situation now in the South China Sea, Crawford says. At the moment there is no way to ensure that scientifically interesting regions remain pristine. We should confront these conundrums before exploitation begins in earnest, says Tony Milligan, an ethicist at King's College London and Elvis's co-author. "Once you have a regular presence on the moon, the law begins to look very different, and the colossal loopholes that you can drive spaceships through suddenly stand out in much sharper relief," he observes.

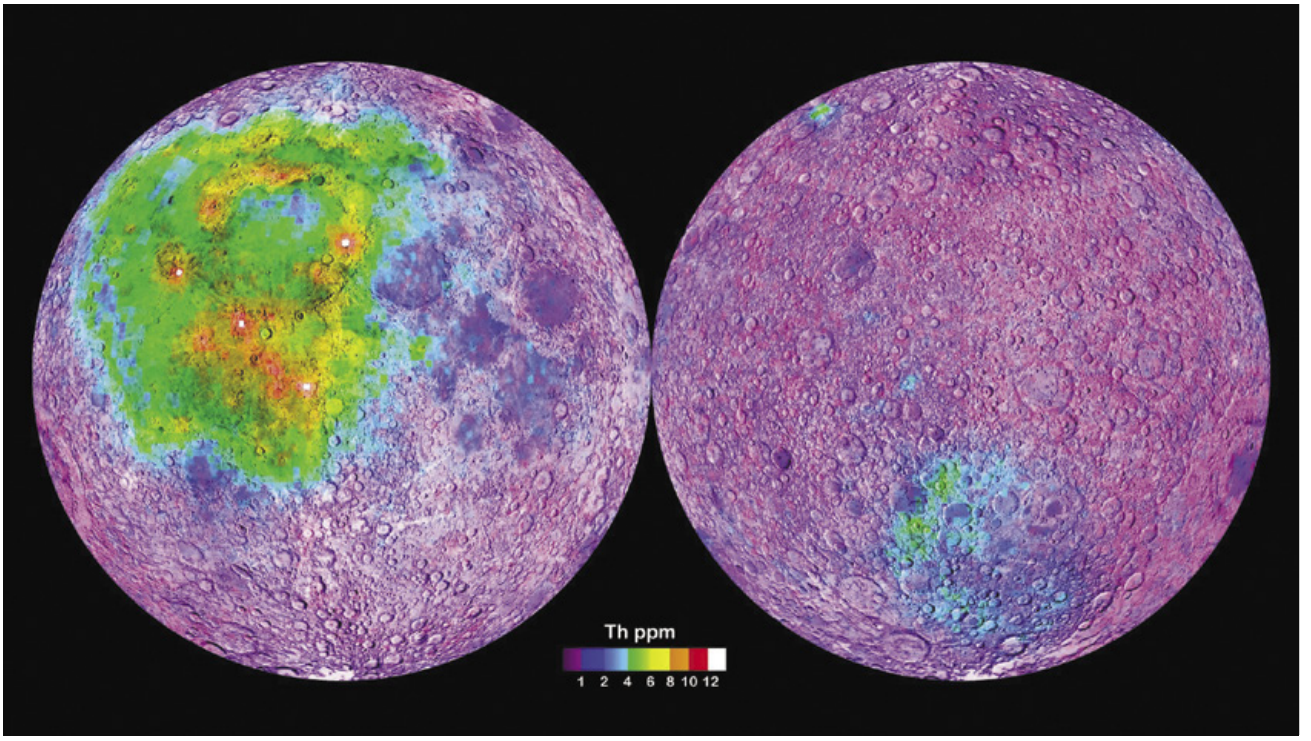
To many, the Chinese space program, with its political will and technological capabilities, appears to have a leg up on the competition. Chinese engineers have suggested that they can place a craft on the surface of the moon with centimeter-scale accuracy. Their upcoming set of Chang'e lunar missions intends to bring back samples and survey the poles in high detail. It might be in the best interests of other countries to begin working on rules that could rein in potential rivals, even if it means giving up some autonomy, von der Dunk says.

Some groups are already trying. A few years ago Tanja Masson-Zwaan, a space law expert at Leiden University in the Netherlands, co-founded the Hague International Space Resources Governance Working Group, an organization that has brought together government, industry and academia, among others, to come up with recommendations for off-world mining. In 2017 the group produced the building blocks for a legal framework of principles that aim to balance the interests of various stakeholders in accordance with international law. Masson-Zwaan recommends establishing something akin to the International Telecommunication Union, an agency at the U.N. that allocates satellite orbits and slices of the radio spectrum among nations for moon mining.

RISKS AND REWARDS

ALTHOUGH PLANS are afoot to grab lunar real estate, extracting resources such as rare elements and shipping them back to Earth is a long-term goal. Some would even say it is a fantasy. Given the incredible cost and technological hurdles involved in simply getting to space, let alone landing on the lunar surface, it is hard to imagine that transporting materials back to Earth would be profitable any time soon. Deep Space Industries and Planetary Resources, two private businesses that were set up in the early part of the 21st century to pursue asteroid mining, both failed to attract enough investment to attempt any deep-space resource extraction; they were eventually acquired by a satellite manufacturer and a cryptocurrency company, respectively.

© 2019 Scientific American



NASA

DATA FROM NASA'S Lunar Prospector mission show the concentration of thorium on the lunar surface. This metal has been suggested as a possible fuel for use in nuclear reactors, making it a potential target for mining.

"It's different than the gold rush days, when anybody with a mule or a pickax could go and try to find gold," says George Sowers, a space resources expert at the Colorado School of Mines.

Still, if economic activities in space take off, mining could follow, experts say. Elvis points to private rocket companies, such as Elon Musk's SpaceX, that are driving down the cost of launching vehicles into orbit. The cheaper and easier it is to get to space, the more common missions will become. Demand for fuel and other resources could follow, and launching materials from the relatively low gravity of the moon would be more cost-effective than from the deep gravitational well of our own planet.

Richards is eager to play his part in bringing the moon within Earth's economic sphere. Yet so far even he has struggled to get above the atmosphere. When Richards co-founded Moon Express in 2010, the company was one of 16 teams competing for the Google Lunar XPRIZE, which challenged a privately funded robotic spacecraft to land on the moon, drive around, and send back pictures and data. The original 2012 deadline for the full grand prize was extended several times, ultimately to March 2018, but in January of that year the XPRIZE Foundation admitted that no one would be able to claim the $30-million purse.

Moon Express now plans to send its first vehicle into lunar orbit in 2020. It remains to be seen, however, whether its business model—offering space agencies and private companies payload rides to the moon—will be viable in the long run. When asked if he sees any conflict between his desire to stake a claim and the need for an equitable solution for everyone, Richards turns philosophical. The tensions in the Outer Space Treaty reflect the tensions between the belief systems of the two countries that wrote it, he says. The Communist Soviets saw the world from a collective perspective in which goods should be equally distributed, whereas the capitalist Americans believed in greater personal freedom and an unfettered private sector. "That's why the treaty is open to interpretation," he says. "I think we have a chance as a species to conquer these new frontiers without having to conquer each other." SA

Adam Mann is a journalist specializing in astronomy and physics. His work has appeared in *National Geographic*, the *Wall Street Journal*, *Wired* and elsewhere.

MORE TO EXPLORE

Forming a Moon with an Earth-like Composition via a Giant Impact. Robin M. Canup in *Science*, Vol. 338, pages 1052–1055; November 23, 2012.

Tidal Evolution of the Moon from a High-Obliquity, High-Angular-Momentum Earth. Matija Ćuk et al. in *Nature*, Vol. 539, pages 402–406; November 17, 2016.

The Origin of the Moon within a Terrestrial Synestia. Simon J. Lock et al. in *JGR Planets*, Vol. 123, No. 4, pages 910–951; April 2018.

Apollo Next Generation Sample Analysis Program: **https://sservi.nasa.gov/articles/apollo-next-generation-sample-analysis-program**

Moon Village Association: **https://moonvillageassociation.org**

Transcripts from *Apollo 11*: **www.hq.nasa.gov/alsj/a11/a11trans.html**

Outer Space Treaty of 1967: **www.state.gov/t/isn/5181.htm**

scientificamerican.com/magazine/sa

© 2019 Scientific American

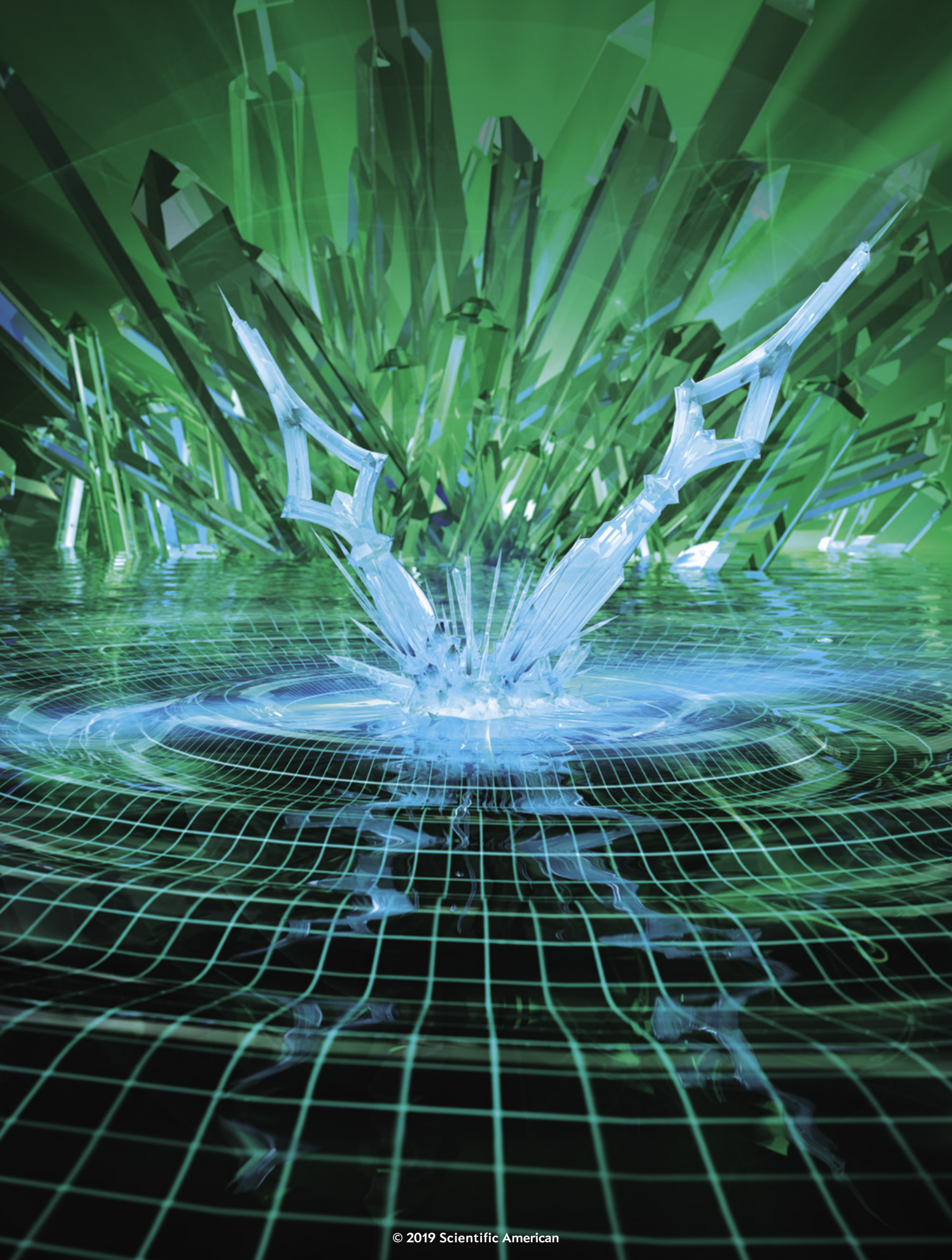
CRYSTALS IN TIME

Surprising new states of matter called time crystals show the same symmetry properties in time that ordinary crystals do in space

By Frank Wilczek

Illustration by Mark Ross Studio

© 2019 Scientific American



© 2019 Scientific American

CRYSTALS ARE NATURE'S MOST ORDERLY SUBSTANCES.

Inside them, atoms and molecules are arranged in regular, repeating structures, giving rise to solids that are stable and rigid—and often beautiful to behold.

People have found crystals fascinating and attractive since before the dawn of modern science, often prizing them as jewels. In the 19th century scientists' quest to classify forms of crystals and understand their effect on light catalyzed important progress in mathematics and physics. Then, in the 20th century, study of the fundamental quantum mechanics of electrons in crystals led directly to modern semiconductor electronics and eventually to smartphones and the Internet.

IN BRIEF

Crystals are orderly states of matter in which the arrangements of atoms take on repeating patterns. In the language of physics, they are said to have "spontaneously broken spatial symmetry."

Time crystals, a newer concept, are states of matter whose patterns repeat at set intervals of time rather than space. They are systems in which time symmetry is spontaneously broken.

The notion of time crystals was first proposed in 2012, and in 2017 scientists discovered the first new materials that fully fit this category. These and others that followed offer promise for the creation of clocks more accurate than ever before.

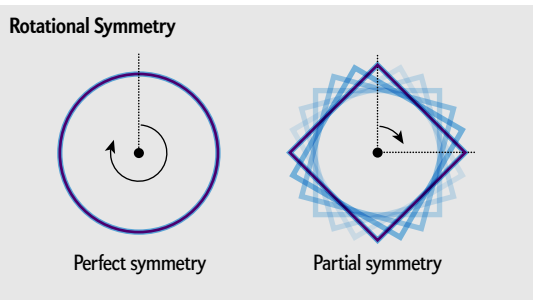
The next step in our understanding of crystals is occurring now, thanks to a principle that arose from Albert Einstein's relativity theory: space and time are intimately connected and ultimately on the same footing. Thus, it is natural to wonder whether any objects display properties in time that are analogous to the properties of ordinary crystals in space. In exploring that question, we discovered "time crystals." This new concept, along with the growing class of novel materials that fit within it, has led to exciting insights about physics, as well as the potential for novel applications, including clocks more accurate than any that exist now.

SYMMETRY

BEFORE I FULLY EXPLAIN this new idea, I must clarify what, exactly, a crystal is. The most fruitful answer for scientific purposes brings in two profound concepts: symmetry and spontaneous symmetry breaking.

In common usage, "symmetry" very broadly indicates balance, harmony or even justice. In physics and mathematics, the meaning is more precise. We say that an object is symmetric or has symmetry if there are transformations that could change it but do not.

That definition might seem strange and abstract at first, so let us focus on a simple example: Consider a circle. When we rotate a circle around its center, through any angle, it remains visually the same, even though every point on it may have moved—it has perfect rotational symmetry. A square has some symmetry but less than a circle because you must rotate a square through a full 90 degrees before it regains its initial appearance. These examples show that the mathematical concept of symmetry captures an essential aspect of its common meaning while adding the virtue of precision.



A second virtue of this concept of symmetry is that it can be generalized. We can adapt the idea so that it applies not just to shapes but more widely to physical laws. We say a law has symmetry if we can change the context in which the law is applied without changing the law itself. For example, the basic axiom of special relativity is that the same physical laws apply when we view the world from different platforms that move at constant velocities relative to one another. Thus, relativity demands that physical laws display a kind of symmetry—namely, symmetry under the platform-changing transformations that physicists call "boosts."

A different class of transformations is important for crystals, including time crystals. They are the very simple yet profoundly important transformations known as "translations." Whereas relativity says the

Illustrations by Jen Christiansen

© 2019 Scientific American

same laws apply for observers on moving platforms, spatial translation symmetry says the same laws apply for observers on platforms in different places. If you move—or "translate"—your laboratory from one place to another, you will find that the same laws hold in the new place. Spatial translation symmetry, in other words, asserts that the laws we discover *anywhere* apply *everywhere*.

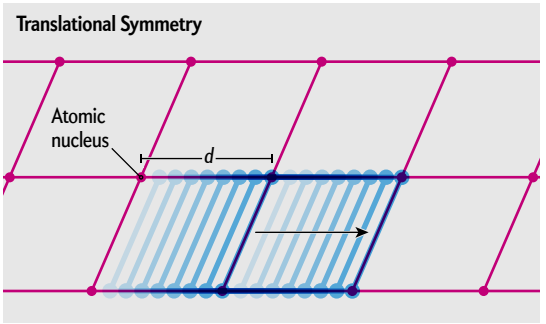
Time translation symmetry expresses a similar idea but for time instead of space. It says the same laws we operate under now also apply for observers in the past or in the future. In other words, the laws we discover at any time apply at every time. In view of its basic importance, time translation symmetry deserves to have a less forbidding name, with fewer than seven syllables. Here I will call it tau, denoted by the Greek symbol τ.

Without space and time translation symmetry, experiments carried out in different places and at different times would not be reproducible. In their everyday work, scientists take those symmetries for granted. Indeed, science as we know it would be impossible without them. But it is important to emphasize that we can test space and time translation symmetry empirically. Specifically, we can observe behavior in distant astronomical objects. Such objects are situated, obviously, in different places, and thanks to the finite speed of light we can observe in the present how they behaved in the past. Astronomers have determined, in great detail and with high accuracy, that the same laws do in fact apply.

SYMMETRY BREAKING

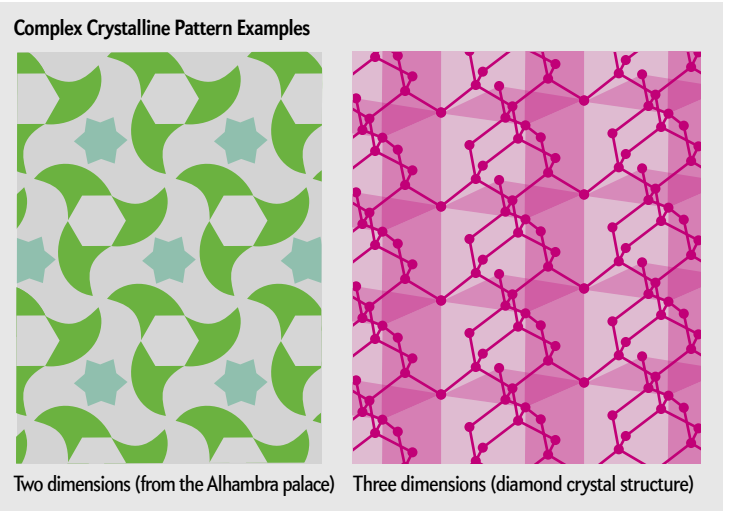
FOR ALL THEIR AESTHETIC SYMMETRY, it is actually the way crystals lack symmetry that is, for physicists, their defining characteristic.

Consider a drastically idealized crystal. It will be one-dimensional, and its atomic nuclei will be located at regular intervals along a line, separated by the distance d. (Their coordinates therefore will be nd, where n is a whole number.) If we translate this crystal to the right by a tiny distance, it will not look like the same object. Only after we translate through the specific distance d will we see the same crystal. Thus, our idealized crystal has a reduced degree of spatial translation symmetry, similarly to how a square has a reduced degree of rotation symmetry.



Physicists say that in a crystal the translation symmetry of the fundamental laws is "broken," leading to a lesser translation symmetry. That remaining symmetry conveys the essence of our crystal. Indeed, if we know that a crystal's symmetry involves translations through multiples of the distance d, then we know where to place its atoms relative to one another.

Crystalline patterns in two and three dimensions can be more complicated, and they come in many varieties. They can display partial rotational and partial translational symmetry. The 14th-century artists who decorated the Alhambra palace in Granada, Spain, discovered many possible forms of two-dimensional crystals by intuition and experimentation, and mathematicians in the 19th century classified the possible forms of three-dimensional crystals.


Complex Crystalline Pattern Examples
Two dimensions (from the Alhambra palace)
Three dimensions (diamond crystal structure)

In the summer of 2011 I was preparing to teach this elegant chapter of mathematics as part of a course on the uses of symmetry in physics. I always try to take a fresh look at material I will be teaching and, if possible, add something new. It occurred to me then that one could extend the classification of possible crystalline patterns in three-dimensional space to crystalline patterns in four-dimensional spacetime.

When I mentioned this mathematical line of investigation to Alfred Shapere, my former student turned valued colleague, who is now at the University of Kentucky, he urged me to consider two very basic physical questions. They launched me on a surprising scientific adventure:

What real-world systems could crystals in spacetime describe?

Might these patterns lead us to identify distinctive states of matter?

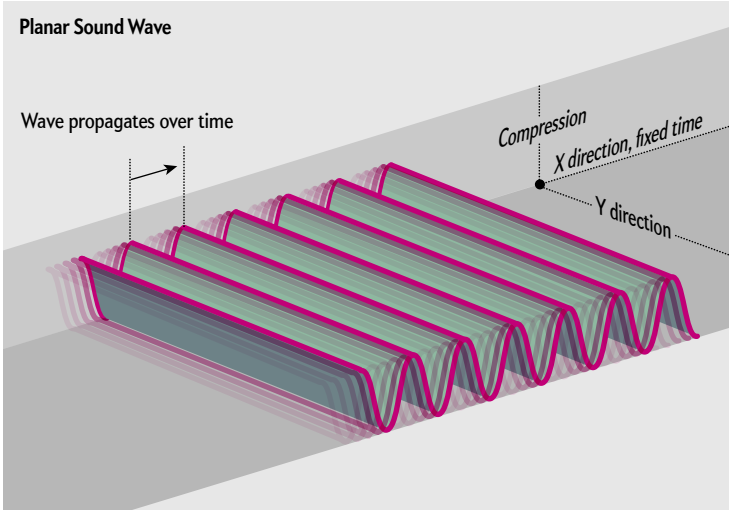
The answer to the first question is fairly straightforward. Whereas ordinary crystals are orderly arrangements of objects in space, spacetime crystals are orderly arrangements of events in spacetime.

As we did for ordinary crystals, we can get our bearings by considering the one-dimensional case, in

© 2019 Scientific American

which spacetime crystals simplify to purely time crystals. We are looking, then, for systems whose overall state repeats itself at regular intervals. Such systems are almost embarrassingly familiar. For example, Earth repeats its orientation in space at daily intervals, and the Earth-sun system repeats its configuration at yearly intervals. Inventors and scientists have, over many decades, developed systems that repeat their arrangements at increasingly accurate intervals for use as clocks. Pendulum and spring clocks were superseded by clocks based on vibrating (traditional) crystals, and those were eventually superseded by clocks based on vibrating atoms. Atomic clocks have achieved extraordinary accuracy, but there are important reasons to improve them further—and time crystals might help, as we will see later.

Some familiar real-world systems also embody higher-dimensional spacetime crystal patterns. For example, the pattern shown here can represent a planar sound wave, where the height of the surface indicates compression as a function of position and time. More elaborate spacetime crystal patterns might be difficult to come by in nature, but they could be interesting targets for artists and engineers—imagine a dynamic Alhambra on steroids.



These types of spacetime crystals, though, simply repackage known phenomena under a different label. We can move into genuinely new territory in physics by considering Shapere's second question. To do that, we must now bring in the idea of *spontaneous* symmetry breaking.

SPONTANEOUS SYMMETRY BREAKING

WHEN A LIQUID or gas cools into a crystal, something fundamentally remarkable occurs: the emergent solution of the laws of physics—the crystal—displays less symmetry than the laws themselves. As this reduction of symmetry is brought on just by a decrease in temperature, without any special outside intervention, we can say that in forming a crystal the material breaks spatial translation symmetry "spontaneously."

An important feature of crystallization is a sharp change in the system's behavior or, in technical language, a sharp phase transition. Above a certain critical temperature (which depends on the system's chemical composition and the ambient pressure), we have a liquid; below it we have a crystal—objects with quite different properties. The transition occurs predictably and is accompanied by the emission of energy (in the form of heat). The fact that a small change in ambient conditions causes a substance to reorganize into a qualitatively distinct material is no less remarkable for being, in the case of water and ice, very familiar.

The rigidity of crystals is another emergent property that distinguishes them from liquids and gases. From a microscopic perspective, rigidity arises because the organized pattern of atoms in a crystal persists over long distances and the crystal resists attempts to disrupt that pattern.

The three features of crystallization that we have just discussed—reduced symmetry, sharp phase transition and rigidity—are deeply related. The basic principle underlying all three is that atoms "want" to form patterns with favorable energy. Different choices of pattern—in the jargon, different phases—can win out under different conditions (for instance, various pressures and temperatures). When conditions change, we often see sharp phase transitions. And because pattern formation requires collective action on the part of the atoms, the winning choice will be enforced over the entire material, which will snap back into its previous state if the chosen pattern is disturbed.

Because spontaneous symmetry breaking unites such a nice package of ideas and powerful implications, I felt it was important to explore the possibility that τ can be broken spontaneously. As I was writing up this idea, I explained it to my wife, Betsy Devine: "It's like a crystal but in time." Drawn in by my excitement, she was curious: "What are you calling it?" "Spontaneous breaking of time translation symmetry," I said. "No way," she countered. "Call it time crystals." Which, naturally, I did. In 2012 I published two papers, one co-authored by Shapere, introducing the concept. A time crystal, then, is a system in which τ is spontaneously broken.

One might wonder why it took so long for the concepts of τ and spontaneous symmetry breaking to come together, given that separately they have been understood for many years. It is because τ differs from other symmetries in a crucial way that makes the question of its possible spontaneous breaking much subtler. The difference arises because of a profound theorem proved by mathematician Emmy Noether in 1915. Noether's theorem makes a connection between symmetry principles and conservation laws—it shows that for every form of symmetry, there is a corresponding quantity that is conserved. In the application relevant here, Noether's theorem states that τ is basically equivalent to the conservation of energy.

© 2019 Scientific American

Conversely, when a system breaks τ, energy is not conserved, and it ceases to be a useful characteristic of that system. (More precisely: without τ, you can no longer obtain an energylike, time-independent quantity by summing up contributions from the system's parts.)

The usual explanation for why spontaneous symmetry breaking occurs is that it can be favorable energetically. If the lowest-energy state breaks *spatial* symmetry and the energy of the system is conserved, then the broken symmetry state, once entered, will persist. That is how scientists account for ordinary crystallization, for example.

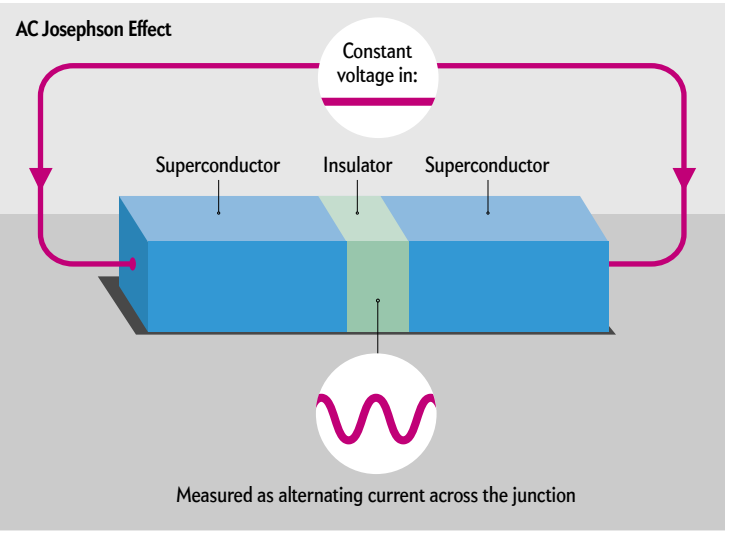
But that energy-based explanation will not work for τ breaking, because τ breaking removes the applicable measure of energy. This apparent difficulty put the possibility of spontaneous τ breaking, and the associated concept of time crystals, beyond the conceptual horizon of most physicists.

There is, however, a more general road to spontaneous symmetry breaking, which also applies to τ breaking. Rather than spontaneously reorganizing to a lower-energy state, a material might reorganize to a state that is more stable for other reasons. For instance, ordered patterns that extend over large stretches of space or time and involve many particles are difficult to unravel because most disrupting forces act on small, local scales. Thus, a material might achieve greater stability by taking on a new pattern that occurs over a larger scale than in its previous state.

Ultimately, of course, no ordinary state of matter can maintain itself against all disruptions. Consider, for example, diamonds. A legendary ad campaign popularized the slogan "a diamond is forever." But in the right atmosphere, if the temperature is hot enough, a diamond will burn into inglorious ash. More basically, diamonds are not a stable state of carbon at ordinary temperatures and atmospheric pressure. They are created at much higher pressures and, once formed, will survive for a very long time at ordinary pressures. But physicists calculate that if you wait long enough, your diamond will turn into graphite. Even less likely, but still possible, a quantum fluctuation can turn your diamond into a tiny black hole. It is also possible that the decay of a diamond's protons will slowly erode it. In practice, what we mean by a "state of matter" (such as diamond) is an organization of a substance that has a useful degree of stability against a significant range of external changes.

OLD AND NEW TIME CRYSTALS

THE AC JOSEPHSON EFFECT is one of the gems of physics, and it supplies the prototype for one large family of time crystals. It occurs when we apply a constant voltage V (a difference in potential energy) across an insulating junction separating two superconducting materials (a so-called Josephson junction, named after physicist Brian Josephson). In this situation, one observes that an alternating current at frequency $2eV/\hbar$ flows across the junction, where e is the charge of an electron and $\hbar$ is the reduced Planck's constant. Here, although the physical setup does not vary in time (in other words, it respects τ), the resulting behavior does vary in time. Full time translation symmetry has been reduced to symmetry under time translation by multiples of the period $\hbar/2eV$. Thus, the AC Josephson effect embodies the most basic concept of a time crystal. In some respects, however, it is not ideal. To maintain the voltage, one must somehow close the circuit and supply a battery. But AC circuits tend to dissipate heat, and batteries run down. Moreover, oscillating currents tend to radiate electromagnetic waves. For all these reasons, Josephson junctions are not ideally stable.



By using various refinements (such as fully superconducting circuits, excellent capacitors in place of ordinary batteries and enclosures to trap radiation), it is possible to substantially reduce the levels of those effects. And other systems that involve superfluids or magnets in place of superconductors exhibit analogous effects while minimizing those problems. In very recent work, Nikolay Prokof'ev and Boris Svistunov have proposed extremely clean examples involving two interpenetrating superfluids.

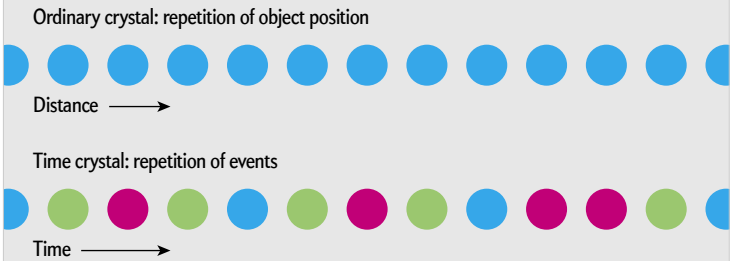
Thinking explicitly about τ breaking has focused attention on these issues and led to the discovery of new examples and fruitful experiments. Still, because the central physical idea is already implicit in Josephson's work of 1962, it seems appropriate to refer to all these as "old" time crystals.

"New" time crystals arrived with the March 9, 2017, issue of *Nature,* which featured gorgeous (metaphorical) time crystals on the cover and announced "Time crystals: First observations of exotic new state of matter." Inside were two independent discovery papers. In one experiment, a group led by Christopher Monroe of the University of Maryland, College Park, created a time crystal in an engineered system of a chain of ytterbium ions. In the other, Mikhail Lukin's group at Harvard University realized a time crystal in a system

© 2019 Scientific American

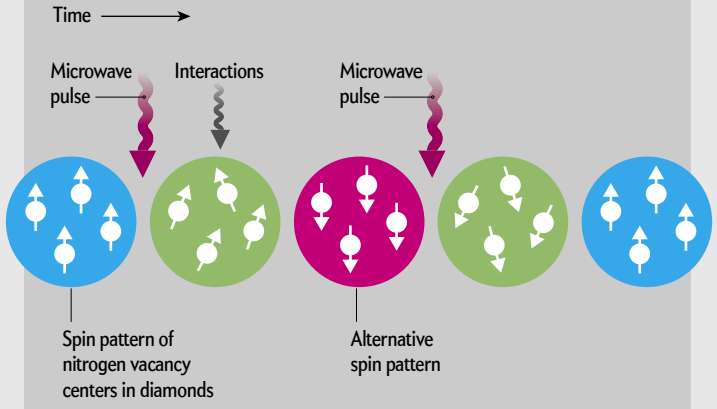
Making a Time Crystal

Just as the atoms in regular crystals repeat their arrangements over certain distances, time crystals are states of matter that repeat over specific periods of time. The first new materials that fit into this category were discovered in 2017 by two research teams, one led by Mikhail Lukin of Harvard University and the other by Christopher Monroe of the University of Maryland, College Park.



The Lukin Experiment

Lukin's group created a time crystal by manipulating the spins of atoms in so-called nitrogen vacancy centers—impurities in a diamond lattice. The researchers periodically exposed the diamond to laser pulses. Between pulses, the spins continued to interact with one another. The entire system repeated its overall configuration periodically—but not with the same period as the microwave pulses. Rather the system took on its own timing period, cycling at a fraction of the frequency of the pulses.



of many thousands of defects, called nitrogen vacancy centers, in a diamond.

In both systems, the spin direction of the atoms (either the ytterbium ions or the diamond defects) changes with regularity, and the atoms periodically come back into their original configurations. In Monroe's experiment, researchers used lasers to flip the ions' spins and to correlate the spins into connected, "entangled" states. As a result, though, the ions' spins began to oscillate at only half the rate of the laser pulses. In Lukin's project, the scientists used microwave pulses to flip the diamond defects' spins. They observed time crystals with twice and three times the pulse spacing. In all these experiments, the materials received external stimulation—lasers or microwave pulses—but they displayed a different period than that of their stimuli. In other words, they broke time symmetry spontaneously.

These experiments inaugurated a direction in materials physics that has grown into a minor industry. More materials utilizing the same general principles—which have come to be called Floquet time crystals—have come on the scene since then, and many more are being investigated.

Floquet time crystals are distinct in important ways from related phenomena discovered much earlier. Notably, in 1831 Michael Faraday found that when he shook a pool of mercury vertically with period T, the resulting flow often displayed period $2T$. But the symmetry breaking in Faraday's system—and in many other systems studied in the intervening years prior to 2017—does not allow a clean separation between the material and the drive (in this case, the act of shaking), and it does not display the hallmarks of spontaneous symmetry breaking. The drive never ceases to pump energy (or, more accurately, entropy), which is radiated as heat, into the material.

In effect, the entire system consisting of material plus drive—whose behavior, as noted, cannot be cleanly separated—simply has less symmetry than the drive considered separately. In the 2017 systems, in contrast, after a brief settling-down period, the material falls into a steady state in which it no longer exchanges energy or entropy with the drive. The difference is subtle but physically crucial. The new Floquet time crystals represent distinct phases of matter, and they display the hallmarks of spontaneous symmetry breaking, whereas the earlier examples, though extremely interesting in their own right, do not.

Likewise, Earth's rotation and its revolution around the sun are not time crystals in this sense. Their impressive degree of stability is enforced by the approximate conservation of energy and angular momentum. They do not have the lowest possible values of those quantities, so the preceding energetic argument for stability does not apply; they also do not involve long-range patterns. But precisely because of the enormous value of energy and angular momentum in these systems, it takes either a big disturbance or small disturbances acting over a long time to significantly change them. Indeed, effects that include the tides, the gravitational influence of other planets and even the evolution of the sun do slightly alter those astronomical systems. The associated measures of time such as "day" and "year" are, notoriously, subject to occasional correction.

In contrast, these new time crystals display strong rigidity and stability in their patterns—a feature that offers a way of dividing up time very accurately, which could be the key to advanced clocks. Modern atomic clocks are marvels of accuracy, but they lack the guaranteed long-term stability of time crystals. More accurate, less cumbersome clocks based on these emerging states of matter could empower exquisite measurements of

© 2019 Scientific American

distances and times, with applications from improved GPS to new ways of detecting underground caves and mineral deposits through their influence on gravity or even gravitational waves. DARPA—the Defense Advanced Research Projects Agency—is funding research on time crystals with such possibilities in mind.

THE TAO OF τ

THE CIRCLE OF IDEAS and experiments around time crystals and spontaneous τ breaking represents a subject in its infancy. There are many open questions and fronts for growth. One ongoing task is to expand the census of physical time crystals to include larger and more convenient examples and to embody a wider variety of spacetime patterns, by both designing new time crystal materials and discovering them in nature. Physicists are also interested in studying and understanding the phase transitions that bring matter into and out of these states.

Another task is to examine in detail the physical properties of time crystals (and spacetime crystals, in which space symmetry and τ are both spontaneously broken). Here the example of semiconductor crystals, mentioned earlier, is inspiring. What discoveries will emerge as we study how time crystals modify the behavior of electrons and light moving within them?

Having opened our minds to the possibility of states of matter that involve time, we can consider not only time crystals but also time quasicrystals (materials that are very ordered yet lack repeating patterns), time liquids (materials in which the density of events in time is constant but the period is not) and time glasses (which have a pattern that looks perfectly rigid but actually shows small deviations). Researchers are actively exploring these and other possibilities. Indeed, some forms of time quasicrystals and a kind of time liquid have been identified already.

So far we have considered phases of matter that put τ into play. Let me conclude with two brief comments about τ in cosmology and in black holes.

The steady-state-universe model was a principled attempt to maintain τ in cosmology. In that model, popular in the mid-20th century, astronomers postulated that the state, or appearance, of the universe on large scales is independent of time—in other words, it upholds time symmetry. Although the universe is always expanding, the steady-state model postulated that matter is continuously being created, allowing the average density of the cosmos to stay constant. But the steady-state model did not survive the test of time. Instead astronomers have accumulated overwhelming evidence that the universe was a very different place 13.7 billion years ago, in the immediate aftermath of the big bang, even though the same physical laws applied. In that sense, τ is (perhaps spontaneously) broken by the universe as a whole. Some cosmologists have also suggested that ours is a cyclic universe or that the universe went through a phase of rapid oscillation. These speculations—which, to date, remain just that—bring us close to the circle of ideas around time crystals.

Finally, the equations of general relativity, which embody our best present understanding of spacetime structure, are based on the concept that we can specify a definite distance between any two nearby points. This simple idea, though, is known to break down in at least two extreme conditions: when we extrapolate big bang cosmology to its initial moments and in the central interior of black holes. Elsewhere in physics, breakdown of the equations that describe behavior in a given state of matter is often a signal that the system will undergo a phase transition. Could it be that spacetime itself, under extreme conditions of high pressure, high temperature or rapid change, abandons τ?

It occurred to me that one could extend the classification of possible crystalline patterns in three-dimensional space to crystalline patterns in four-dimensional spacetime.

Ultimately the concept of time crystals offers a chance for progress both theoretically—in terms of understanding cosmology and black holes from another perspective—and practically. The novel forms of time crystals most likely to be revealed in the coming years should move us closer to more perfect clocks, and they may turn out to have other useful properties. In any case, they are simply interesting and present us with opportunities to expand our ideas about how matter can be organized. SA

Frank Wilczek is a theoretical physicist at the Massachusetts Institute of Technology. He won the 2004 Nobel Prize in Physics for his work on the theory of the strong force, and in 2012 he proposed the concept of time crystals.

MORE TO EXPLORE

Classical Time Crystals. Alfred Shapere and Frank Wilczek in *Physical Review Letters,* Vol. 109, No. 16, Article No. 160402; October 2012.
Quantum Time Crystals. Frank Wilczek in *Physical Review Letters,* Vol. 109, No. 16, Article No. 160401; October 2012.
Observation of a Discrete Time Crystal. Jiehang Zhang et al. in *Nature,* Vol. 543, pages 217–220; March 9, 2017.
Observation of Discrete Time-Crystalline Order in a Disordered Dipolar Many-Body System. Soonwon Choi et al. in *Nature,* Vol. 543, pages 221–225; March 9, 2017.
Time Crystals: A Review. Krzysztof Sacha and Jakub Zakrzewski in *Reports on Progress in Physics,* Vol. 81, No. 1, Article No. 016401; January 2018.
Time Crystals in Periodically Driven Systems. Norman Y. Yao and Chetan Nayak in *Physics Today,* Vol. 71, No. 9, pages 40–47; September 2018.

scientificamerican.com/magazine/sa

© 2019 Scientific American

© 2019 Scientific American

The rise of Animals

New fossils and analyses of ancient ocean chemistry reveal the surprisingly deep roots of the Cambrian explosion

By Rachel A. Wood

Illustration by Franz Anthony

© 2019 Scientific American

STAND ATOP THE STEEP WHITE CLIFFS THAT SURROUND the giant rivers of Siberia, and your feet will mark a pivotal point in the history of life on Earth: the 541-million-year-old geologic boundary between the Precambrian and Cambrian periods. The rocks below this dividing line contain scant fossil remains—ghostly impressions of soft-bodied organisms and a smattering of shelly forms. But break open any of the rocks just above the boundary, and they will be teeming with shells. A little higher up still, familiar fossil creatures such as trilobites appear. These changes document the so-called Cambrian explosion, one of the most significant, but still poorly understood, events in all of evolution.

IN BRIEF

Scientists long thought that complex animals originated during the Cambrian explosion. **But mounting fossil evidence** indicates that they actually made their debut millions of years earlier, during the Ediacaran period. **New techniques** for reconstructing ancient ocean chemistry have yielded insights into the environmental pressures that drove this early evolutionary diversification.

For decades scientists thought that complex animals—multicellular organisms with differentiated tissue types—originated in the Cambrian explosion. To be sure, a riot of novel forms burst into existence during this time, including the ancestors of many of today's major animal groups. But recent discoveries from Siberia, Namibia and elsewhere show that complex animals actually got their start millions of years before the Cambrian explosion, during the last chapter of the Precambrian, known as the Ediacaran. Among these finds are the oldest known creatures with external and internal skeletons composed of mineralized tissue, a pivotal evolutionary innovation seen in many modern-day animals.

The presence of these armored creatures so far back in time—550 million years ago—indicates that the ecological and environmental pressures thought to have driven the Cambrian explosion were in fact at work long before then. Figuring out how these factors shaped the evolution of the earliest complex animals in the Ediacaran is key to understanding the astonishing burst of diversification that followed in the Cambrian.

The Cambrian fossil record has been the subject of intense study for more than 150 years. Thus, the broad global patterns of what Cambrian fossils appeared when—and where—are relatively well established: similar fossils turned up on many continents at around the same time, and they followed the same succession of evolutionary changes more or less synchronously. But only now, with the discoveries of the older Ediacaran fossils, are we starting to see the roots of the Cambrian explosion.

Gratifyingly, we are also beginning to puzzle out why it happened when it did, thanks in part to the development of new geochemical techniques that have revolutionized our understanding of the changing chemistry of the oceans in the Ediacaran-Cambrian world. Insights from the emerging fossil and geochemical records have just recently been integrated to show how the planet's biosphere, geosphere, hydrosphere and atmosphere—together known as the Earth system—may have operated during this interval. But already we can paint a striking picture of how the seafloor became successively populated by ever more complex creatures tens of millions of years before the Cambrian explosion, setting the stage for the rise of animal life as we know it.

© 2019 Scientific American

THE FIRST ANIMALS

THE OLDEST POSSIBLE EVIDENCE of ancient animals comes not from recognizable fossils but from the remains of organic compounds known as biomarkers. Researchers have found one such biomarker, a particular form of sterane, in very well preserved rocks from a sedimentary sequence known as the Huqf Supergroup in Oman, which is at least 650 million years old. Some experts have argued that these steranes are unique to a particular group of sponges and that the presence of the molecules in the Huqf rocks therefore documents the existence of these animals at that very early time. Not all scientists accept the assertion that these steranes are specific to those sponges, however. Indeed, a study published in April suggests that they are diagnostic of a group of single-celled forms of amoebas.

The oldest candidate animal fossils, which hail from a sequence of rocks in southern China called the Lantian Formation and are possibly as old as 635 million years, are similarly contested. Some investigators think these tiny, soft-bodied forms are related to corals or jellyfish because they exhibit tentaclelike structures, but the preservation of these fossils is not sufficiently clear to allow unequivocal interpretation, leaving many researchers unconvinced that they represent animals of any kind.

The oldest animal remains that almost everyone can agree on are fossils from Newfoundland that date to about 571 million years ago, shortly after the last regional "Snowball Earth" glaciation that encased much of the planet in thick ice. These earliest known representatives of the Ediacaran biota were dominated by soft-bodied creatures up to a meter in height or width. Some took the form of large, featherlike fronds with vertical stalks that rooted them to the seafloor; others sprawled across the ocean bottom, their flat bodies exhibiting a fractal architecture, with branching units that showed the same patterns at all scales. All these body plans maximize surface area, suggesting that these animals absorbed nutrients directly from the surrounding water.

This modest variety of fauna prevailed for more than 10 million years. But then the pace of animal evolution began to accelerate. The fossil record indicates that after around 560 million years ago, the Ediacaran biota diversified to include mobile forms that inhabited shallow seas. Some of the fossils preserve scratch marks that suggest the animals were eating algal mats by grazing. Others may have dragged themselves across the algae, absorbing nutrients from the underside of their bodies. The first simple burrows also appear at around this time, evidence that animals had started to move and disturb the sediment of the seafloor.

Fast-forward to around 550 million years ago, and the oldest fossils preserving external and internal skeletons suddenly appear in limestone rocks (which consist mainly of calcium carbonate). These fossils are already diverse in size and form, and they show up in such far-flung locales as Siberia, Brazil and Namibia. The presence of skeletons in so many unrelated animal groups around the world at this point in time is testament to a major driving evolutionary force operating on a global scale. We do not know for sure what this force was. But we have an idea. Making a skeleton is energetically expensive, so for an animal to undertake such an endeavor the benefit must outweigh the cost. Animals may produce a skeleton for many reasons, but by far the most common is the need for protection from predators. Although there is no fossil evidence of predators from this time period, it stands to reason that the appearance of skeletons might reflect the first widespread occurrence of animals that ate other animals.

FOSSIL TRAIL: Key fossils of complex animals have come from Ediacaran rocks on the shores of the Yudoma River in Siberia (*1*) and on the edge of the Nama Desert in Namibia (*2*).

STRONGER TOGETHER

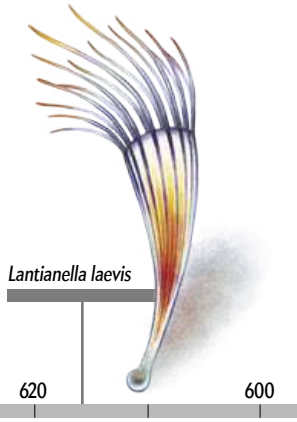
RECENT ANALYSES of these ancient skeletons have yielded tantalizing clues to what their owners looked like—and how they lived. Known from fossils of its delicate tubular skeleton that grew up to about 70 millimeters long and resembled a stack of ice cream cones, an organism called *Cloudina* has figured importantly in our reconstructions of Ediacaran ecosystems. *Cloudina*

RACHEL A. WOOD

© 2019 Scientific American

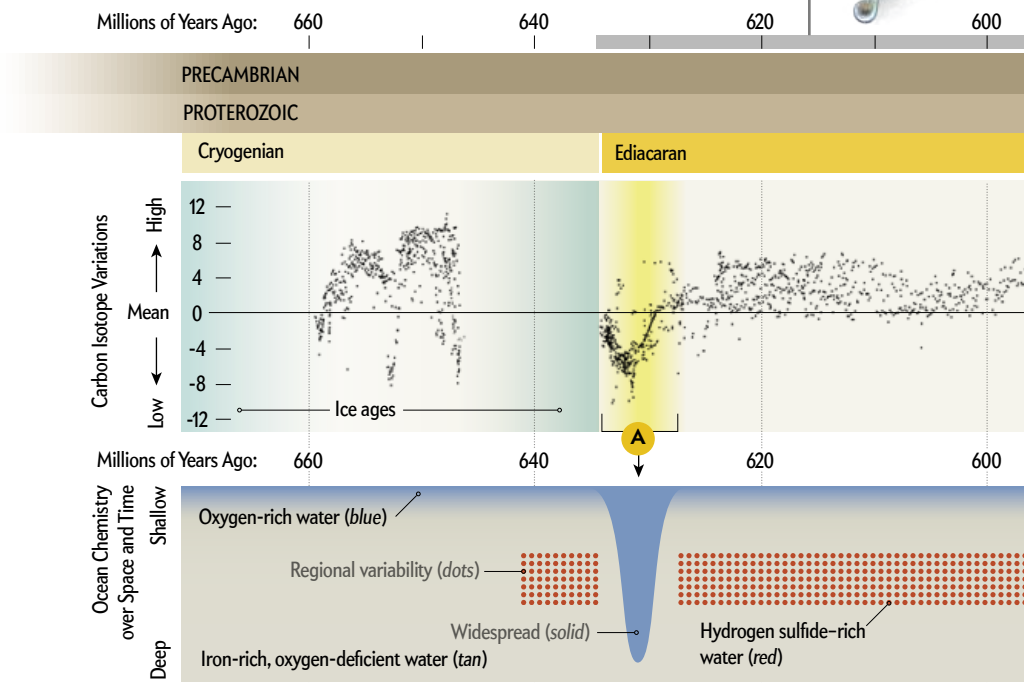
Before the Cambrian Explosion

Many key innovations in animal evolution that were traditionally thought to have originated in the Cambrian actually trace back much further in time to the Ediacaran. For example, the first animals with skeletons debuted during this earlier period. Their ability to produce mineralized tissue probably evolved as a means of protection from predators. Integrating the fossil and geochemical records spanning the time between 670 million and 480 million years ago reveals clues to the environmental factors driving this early evolutionary activity.



Geochemical Evidence

Animals need oxygen to survive. The evolutionary diversification that took place during the Ediacaran occurred under wildly fluctuating oxygen levels in the world's oceans. Carbon isotopes from Ediacaran rocks show that the carbon cycle was unstable and in a state of flux. Analyses of the iron compounds in these rocks, meanwhile, show that dissolved oxygen in the oceans probably reached a threshold or a series of thresholds in the Ediacaran that allowed animals to diversify by meeting their increased metabolic demands as they became more active. Researchers now believe the seas became progressively oxygenated not as one slow, gradual increase but in a series of episodes (A, B, C and D) that appear to coincide with the carbon isotope variations. This trend continued throughout the Ediacaran and probably well beyond.



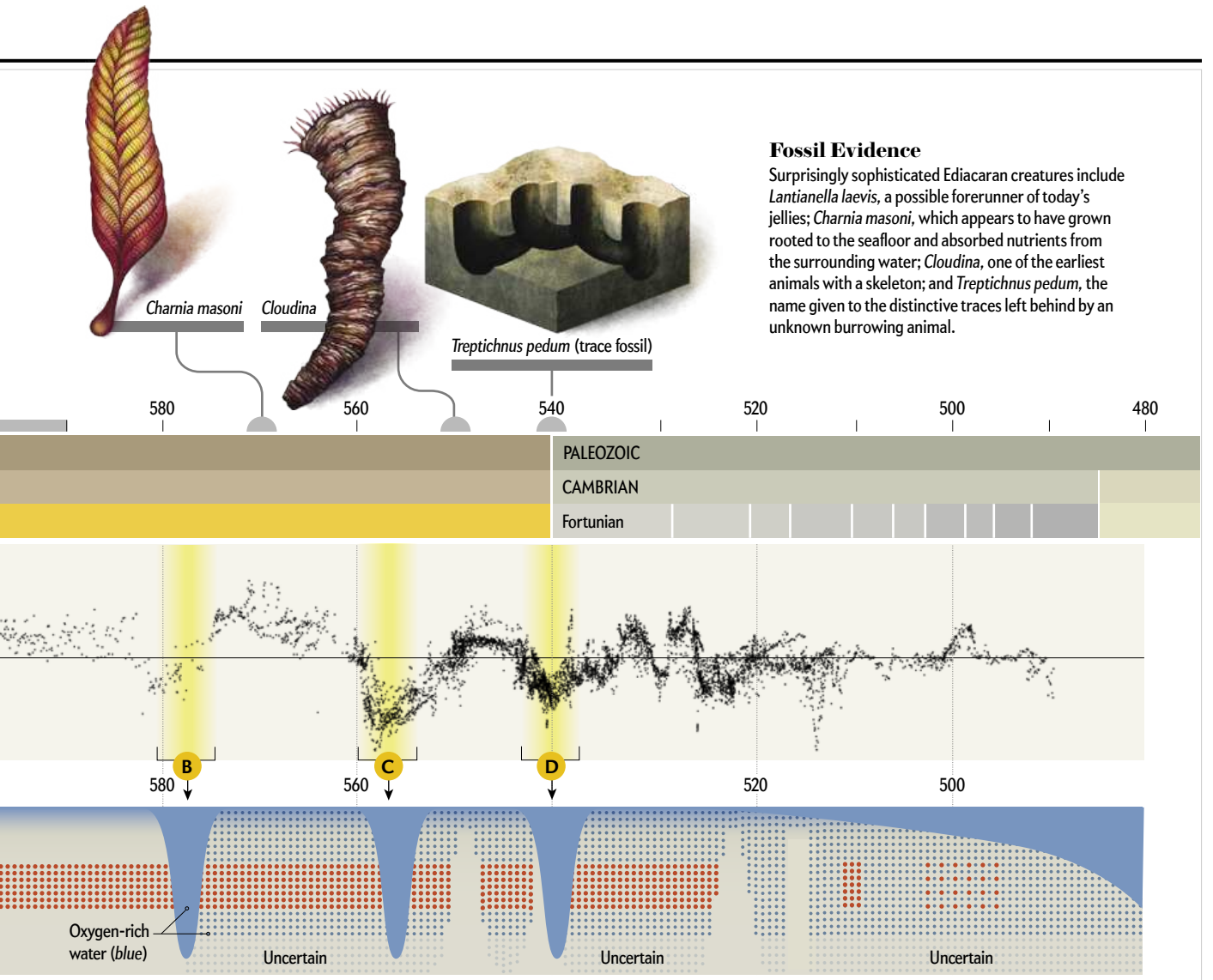
was first discovered in Namibia in 1972 and was long assumed to have grown attached to the seafloor. But in the past few years researchers have identified many new specimens of *Cloudina* from sites around the world that have changed that view. My team's work on specimens from Namibia has shown that *Cloudina* had a variety of growth styles. It could attach to mats made of microbes that bound the soft sediment of the seafloor, or it could anchor itself to layered mounds of cyanobacteria. Most important of all, *Cloudina* individuals could actually cement themselves to one another to form a reef. This finding has established *Cloudina* as one of the oldest reef-building animals, pushing back the record of this way of life by some 20 million years.

Whether *Cloudina* was related to modern reef builders such as corals remains uncertain. But we do know that like reef-building corals, it lived in proximity to a number of other animals. Hints of this intimate association have come from other skeletal fossils found in rocks of the same age as those that contain *Cloudina* fossils. A creature called *Namacalathus,* known from fossil localities around the world, appears to have been one of *Cloudina*'s consorts. Its skeleton was up to 50 millimeters long, composed of a delicate, thin-walled stalk and a cup with a central opening at the top and several openings around the sides. The animal's soft tissue was probably mainly inside the cup, although it is never preserved. Fossils of *Namacalathus* indicate that it grew rooted to microbial mats, often near *Cloudina*.

Namapoikia, a creature known only from fossil localities in Namibia, also fraternized with *Cloudina*. This animal is remarkable for its large size—up to one meter in diameter—and robust skeleton. On the basis of its growth form, we think *Namapoikia* was a sponge and so would have had an internal skeleton, in contrast to the external skeletons *Cloudina* and *Namacalathus* probably had. Intriguingly, *Namapoikia* grew within the hidden places of reefs, encrusting the vertical walls of open cracks and fissures. In modern reefs

Graphics by Jen Christiansen

© 2019 Scientific American



Fossil Evidence

Surprisingly sophisticated Ediacaran creatures include *Lantianella laevis*, a possible forerunner of today's jellies; *Charnia masoni*, which appears to have grown rooted to the seafloor and absorbed nutrients from the surrounding water; *Cloudina*, one of the earliest animals with a skeleton; and *Treptichnus pedum*, the name given to the distinctive traces left behind by an unknown burrowing animal.

SOURCE: "INTEGRATED RECORDS OF ENVIRONMENTAL CHANGE AND EVOLUTION CHALLENGE THE CAMBRIAN EXPLOSION," BY RACHEL WOOD ET AL., IN *NATURE ECOLOGY & EVOLUTION*, VOL. 3; APRIL 2019; GRAPHIC BY JEN CHRISTIANSEN

the communities of animals and plants that live on open surfaces differ from those that occupy these more hidden areas such as caves, crevices or underhangs. Our Ediacaran fossil discoveries indicate that this distinction is as old as animal reefs themselves.

These observations are significant because reef building represents an important ecological innovation. By growing closely together and even cementing to one another, individuals can become mechanically stronger, rise above the seafloor away from competitors, enhance feeding efficiency and gain protection from predators. Like the earliest skeletons, then, the appearance of reefs in the Ediacaran fossil record may signal rising, complex ecological pressures. The Cambrian explosion, and indeed an arms race between predator and prey, had already begun.

THE EDIACARAN WORLD

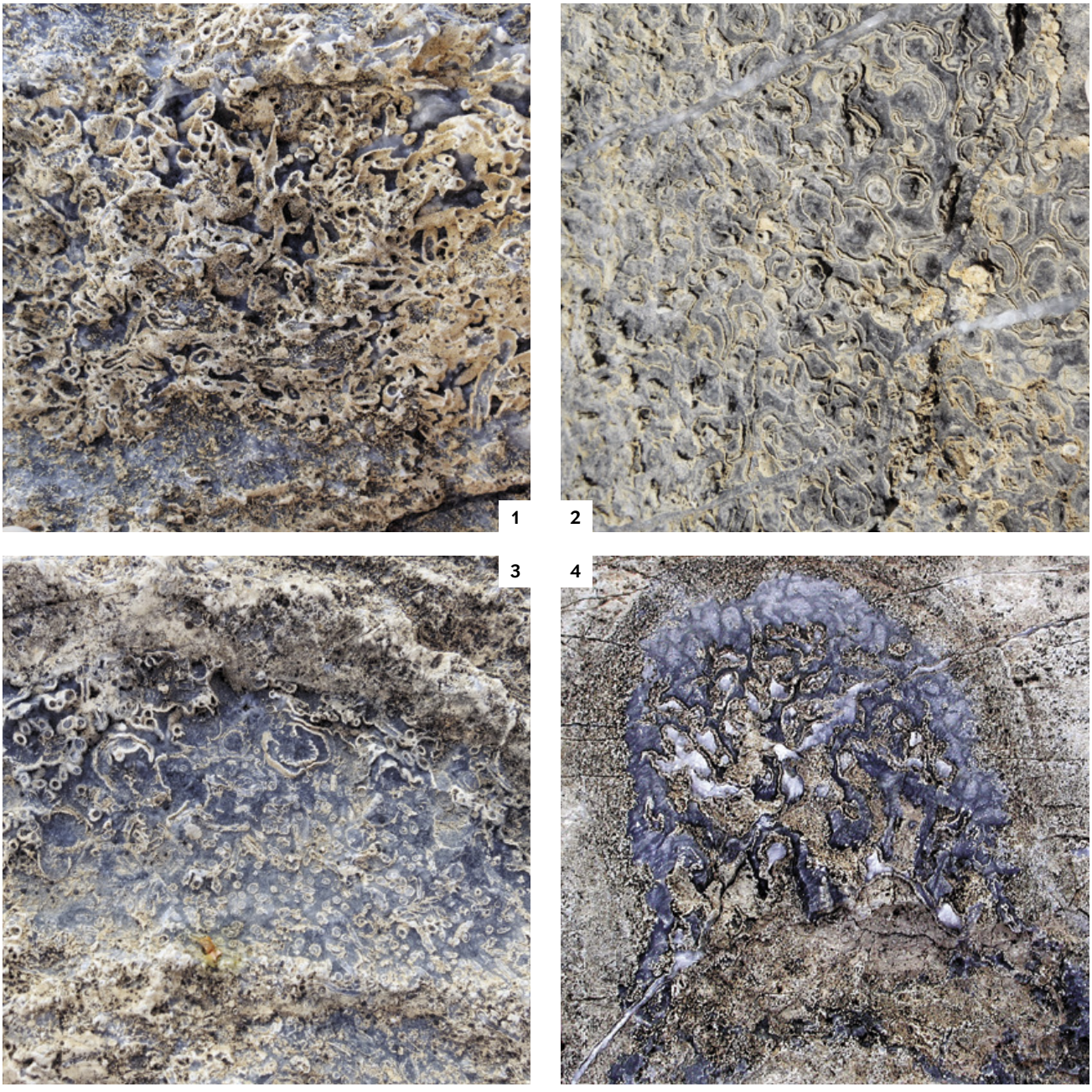
BY THE MID-2010S it was becoming clear that the Cambrian did not mark the sudden, dramatic departure from the Ediacaran that experts long envisioned. Not only had researchers begun to amass evidence that animals started evolving skeletons and building reefs earlier than traditionally thought, but we had also developed ecosystem models showing that Ediacaran animal communities shared many ecological traits with Cambrian ones. The "explosion," we were learning, had a far longer fuse than was previously recognized.

Then, a few years ago, some key discoveries in Siberia and China blurred the Ediacaran and Cambrian worlds even further. A group of researchers from China and Germany found that *Cloudina* persisted into the Cambrian. And my group, together with colleagues from Russia and China, found fossils long thought to be unique to the Cambrian in Ediacaran rocks. These findings underscored to us that to solve the mystery of the Cambrian explosion we had to figure out the dynamics of the Ediacaran world in which these animals originated.

The possible role of shifting oxygen availability is

Illustration by Franz Anthony (fossils)

© 2019 Scientific American



WRITTEN IN STONE: *Cloudina* **is one of the oldest organisms with an external skeleton, known from fossils that preserve its delicate, tubular armor (*1*). *Cloudina* individuals could cement themselves to one another, forming reefs. Another early skeletal form is *Namacalathus*, whose cup-shaped hard parts are typically preserved en masse (*2*). *Cloudina* and *Namacalathus* are often found together (*3*). *Namapoikia*, an early sponge, was another of *Cloudina*'s associates, one that grew in the hidden crevices of the reef (*4*).**

one dynamic that researchers have been particularly keen to pin down. Animals need oxygen, so a central debate over the past few years has been to understand whether at some point in the time spanning the Ediacaran and Cambrian oxygen levels rose beyond a certain critical threshold, allowing animals to flourish. The question is more complicated than it might seem because animals do not all have the same oxygen requirements. Simple, immobile creatures, such as sponges, may need less oxygen than mobile animals, and they certainly require far less of the stuff than active, fast-swimming predators do. We have borne this variation in mind in the course of our investigations.

Fortunately for us, many new geochemical methods for estimating how much oxygen existed in these ancient seas have been developed in recent years. One especially powerful technique—Fe speciation—harnesses the characteristics of the various compounds of iron, which behave differently depending on whether oxygen is present or not. This method allows us to see at a local

RACHEL A. WOOD

© 2019 Scientific American

scale where—and when—there was enough oxygen to support complex life. Studies carried out using this approach have led to a broad consensus: dissolved oxygen in the oceans probably reached a threshold or series of thresholds during the Ediacaran that allowed animals to diversify by meeting their increasing metabolic demands as they became more mobile and active.

Scientists have now assembled sufficiently large geochemical data sets that we can reconstruct how oxygen was distributed not just at individual Ediacaran sites of a certain age but globally through time. This work reveals patterns throughout the Ediacaran and early Cambrian that differ considerably from today's, with many areas showing a relatively thin veneer of well-oxygenated shallow waters laying atop a thicker wedge of deeper seawater that probably lacked oxygen altogether, a state known as anoxia.

These geochemical data also show that the boundary between the anoxic and oxic waters was very dynamic during this interval, rising and falling with shifting sea levels. Areas of shallow marine seafloor habitable to early animals were thus even more restricted than scientists expected—veritable oases of oxygenated water. If the evolutionary diversification that took place during the Ediacaran and Cambrian occurred under relatively low oxygen levels but with highly dynamic conditions that fluctuated on ecological, global and evolutionary timescales, how might these factors have shaped that extraordinary radiation?

ENGINE OF INNOVATION?

PERIODS OF INCREASED ANOXIA on the seafloor coincide with some well-known mass extinctions, such as the one that punctuated the Permian period 252 million years ago, killing off more than 90 percent of all marine species. But several major diversifications—including those in the Ediacaran-Cambrian, the Ordovician 100 million years later and the mid-late Triassic about 247 million years ago—began during long intervals of dynamic shallow marine anoxia. Considering these events, my colleague Doug Erwin of the Smithsonian Institution and I hypothesized that fluctuating oxygen conditions may have created critical opportunities for evolutionary innovation in soft-bodied animals.

It is far easier for animals to form a skeleton of limestone—the material that makes up the skeletons and shells of many modern marine creatures—when seawater oxygen levels exceed 10 micromoles per liter. Perhaps soft-bodied animals were only able to evolve these calcium carbonate skeletons once oxygen levels reached such a threshold, allowing formerly isolated oases to expand, connect and achieve stability on a global scale.

Much remains to be discovered about how life might have responded to changes in oxygen availability over evolutionary timescales. The response was probably complicated because animals were also contending with additional factors such as the rise of predation. Moreover, feedbacks among individual organisms, ecosystems and the broader Earth system—which are largely unknown—would have also figured into the equation.

We have our work cut out for us. Dramatic changes in the regional processes that shaped Earth's crust throughout the Ediacaran-Cambrian interval have produced many significant gaps in the geologic and fossil record. This means that we have to piece together our narrative about the rise of complex animals from data collected from a multitude of localities all over the world. The fact that many of the key Ediacaran localities are still poorly dated further complicates our task. We typically date rocks of this age by measuring the ratio of lead to uranium in zircon crystals found in nearby layers of ash from ancient volcanic eruptions. This is one of the few methods that can supply an absolute, radiometric age for a given rock. But frustratingly, many of our best-known successions lack these vital ash beds. As a result, we are unable to accurately correlate evolutionary changes that have occurred in different parts of the world, which is essential for creating a solid timeframe for our history of events. A prime example is China's hotly debated Lantian Formation, which has yielded the oldest candidate animal fossils, but whose age could fall anywhere between 635 million and 590 million years.

Nevertheless, there are reasons for optimism. New ash beds are coming to light, and dating methods are being refined. For instance, the ash beds that many research groups use to calculate the ages of the Ediacaran fossils found in Namibia have recently been redated, and the youngest ones—those nearest the Precambrian-Cambrian boundary—have proved to be more than two million years younger than previously thought. This result raises important questions about how these fossils actually correlate with their counterparts in Newfoundland and Siberia, among other key localities. In addition, geochemists are developing new isotopic techniques and other methods that can bring our picture of oxygen conditions in this ancient world into sharper focus. And my team and others are finding new fossils in remote places that have gone largely unexplored until now, such as Siberia.

Sometime in the not so distant future, when we stand on those cliffs, surveying the vast forest below, we will have a far deeper understanding of this most extraordinary slice of time. SA

Rachel A. Wood is a paleontologist and geologist at the University of Edinburgh. Her research focuses on the origins and paleoecology of reefs and the evolution of seawater chemistry.

MORE TO EXPLORE

Low-Oxygen Waters Limited Habitable Space for Early Animals. R. Tostevin et al. in *Nature Communications,* Vol. 7, Article No. 12818; September 23, 2016.

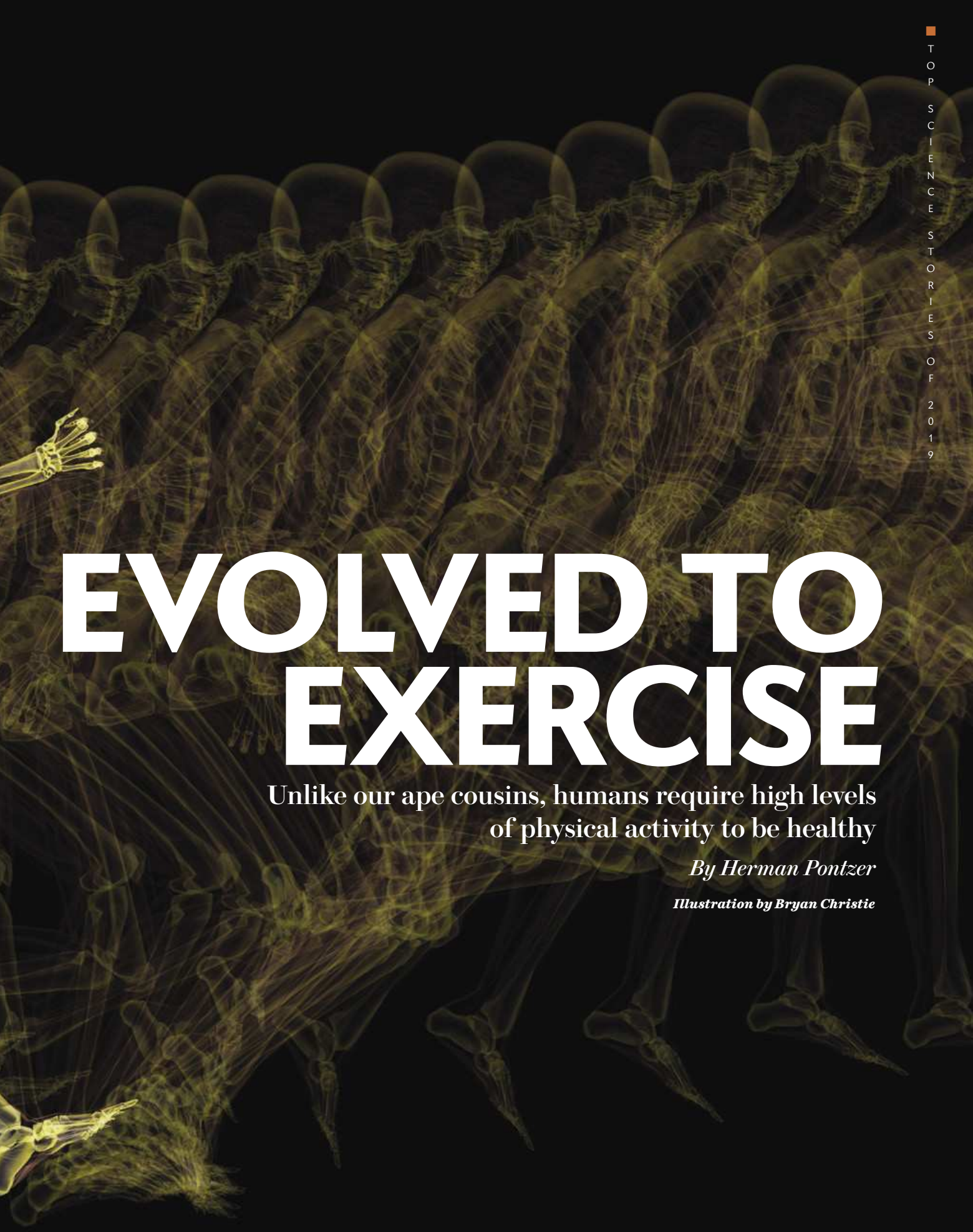
A Deep Root for the Cambrian Explosion: Implications of New Bio- and Chemostratigraphy from the Siberian Platform. M. Zhu et al. in *Geology,* Vol. 45, No. 5, pages 459–462; May 1, 2017.

Integrated Records of Environmental Change and Evolution Challenge the Cambrian Explosion. Rachel Wood et al. in *Nature Ecology & Evolution,* Vol. 3, pages 528–538; April 2019.

scientificamerican.com/magazine/sa

© 2019 Scientific American

© 2019 Scientific American



EVOLVED TO EXERCISE

Unlike our ape cousins, humans require high levels of physical activity to be healthy

By Herman Pontzer

Illustration by Bryan Christie

© 2019 Scientific American

IN THE PREDAWN DAMP OF A UGANDAN RAIN FOREST NEARLY 20 YEARS AGO, I stared up through the crowded canopy at a party of eight chimpanzees sleeping overhead. Our team of three researchers and two field assistants had woken up an hour before, wiggling into rubber boots and hastily assembling backpacks before setting out on muddy trails by headlamp. Now at our destination, the lights were off, and we stood there silently, submerged in a black ocean of forest, the surface 30 meters above, listening to the chimps chuffing and shifting in their leafy nests.

IN BRIEF

Our closest living relatives, the great apes, have habitually low levels of physical activity yet suffer no ill health effects from being lazy.

Humans have evolved to require far higher levels of exercise to be healthy.

New research reveals that as human anatomy and behavior shifted over the past two million years, so, too, did physiology.

Our physiology adapted to the intensive physical activity that hunting and gathering requires.

As a young Ph.D. student studying human and ape evolution, I was in Kibale National Park that summer to measure how much chimpanzees climb each day. It seemed to me that the energy spent climbing might be a critical factor in chimpanzee ecology and evolution, shaping their anatomy to maximize climbing efficiency, thus sparing calories for reproduction and other essential tasks. Months earlier, while mulling over summer research plans from the comfort of my desk at snowy Harvard University, I envisioned chimpanzees waging a heroic struggle for existence, working hard on a daily basis to eke out a living. But as I settled into the rhythm of fieldwork that summer, following chimpanzees from dawn to dusk, I came to a very different conclusion: chimpanzees are lazy. Only recently have I come to appreciate what ape idleness tells us about human evolution.

People are drawn to apes because we see so much of ourselves in them. It is not just that we share more than 97 percent of our DNA with orangutans, gorillas, chimpanzees and bonobos. Apes are clever, use tools, fight and make up, and sneak off to have sex. Some will kill their neighbors over turf and hunt other species for food. The kids learn from their mothers, wrestle and play with one another, and throw tantrums. And the further back in time we go in the fossil record, the more apelike our ancestors look. No species alive today is a perfect model of the past—all lineages change over time. But living apes provide the best chance to see where we came from and to understand how much of us is ancient and unchanged.

And yet it is the differences, rather than the similarities, between humans and apes that are casting new light on the way our bodies work. Discoveries from fossil excavations, zoos and laboratories around the world are revealing just how radically our bodies changed over the past two million years. For decades researchers have known that this last chapter of our evolution was marked by major anatomical and ecological changes—among them, ballooning brain size, hunting and scavenging, increasingly complex stone tools and larger body size. But they have generally assumed that these were changes in shape and behavior, not in the fundamental function of our cells. Current advances are overturning that view, showing how humans have changed physiologically as well. Unlike our ape cousins, we have evolved a dependency on physical activity. We must move to survive.

PARADISE LOST

A TYPICAL DAY'S AGENDA for a chimpanzee in the wild reads like the daily schedule for lethargic retirees on a Caribbean cruise, though with fewer organized activities. Wake up early, crack of dawn, then off to breakfast (fruit). Eat until you are stuffed, and next find a nice place for a nap, maybe some light grooming. After an hour or so (no rush!), go find a sunny tree with figs and gorge yourself. Maybe go meet some friends, a bit more grooming, another nap. Around five o'clock have an early dinner (more fruit, maybe some leaves), then it is time to find a nice sleeping tree, build a nest and call it a night. Sure, there are frenetic pant-hoot choruses when the fruit is really great and the occasional scuffle or monkey hunt, and the alpha male needs to carve some time out every day to thrash a few victims or display mightily. But in general, chimpanzee life is pretty mellow.

It is not just chimps. Orangutans, gorillas and bonobos also lead the sorts of seemingly idle lives that children's fables and high school drug programs warn you about. Great apes spend eight to 10 hours a day resting, grooming and eating before knocking off in the evening for nine or 10 hours of sleep per night.

© 2019 Scientific American

DAVID CAYLESS Getty Images

LOUNGING AROUND: A mountain gorilla family relaxes in Rwanda. Great apes remain healthy at low activity levels.

Chimps and bonobos walk about three kilometers a day, and gorillas and orangutans travel even less. And the climbing? As I discovered that summer, chimpanzees climb about 100 meters a day, the caloric equivalent of another 1.5 kilometers of walking. Orangutans do about the same, and although their ascent has yet to be measured, gorillas surely climb less.

In humans, these activity levels would be a recipe for serious health problems. Our taking fewer than 10,000 daily steps is associated with increased risk of cardiovascular and metabolic disease. U.S. adults typically clock about 5,000 steps, which contributes to the alarming rates of type 2 diabetes, affecting one in 10 Americans, and heart disease, which accounts for a quarter of all deaths in the U.S. By these lights, apes should be in trouble. Converting their walking and climbing to steps per day for comparison across species, we see that great apes rarely accumulate even the modest step counts seen among sedentary humans and never approach the human benchmark of 10,000 steps a day.

Then there is all the sitting and resting. In humans, sitting at a desk or in front of the television for protracted periods is associated with increased risk of disease and a shorter life span, even among people who exercise. Worldwide, physical inactivity is arguably on par with smoking as a health risk, killing more than five million people annually. Among Scottish adults, those watching more than two hours of television a day had a 125 percent increase in cardiac events such as heart attack or stroke. A study in Australian adults reported that every hour accumulated watching television shortened life expectancy by 22 minutes. I will save you the math: bingeing all 63½ hours of *Game of Thrones* in its entirety will cost you one day on this planet.

Yet chimpanzees and other apes remain remarkably healthy at their habitually low levels of physical activity. Even in captivity, diabetes is rare, and blood pressures do not increase with age. Despite having naturally high cholesterol levels, chimpanzee arteries do not harden and clog. As a result, chimps do not develop humanlike heart disease or have heart attacks from occluded coronary arteries. And they stay lean. In 2016 I worked with Steve Ross at Lincoln Park Zoo in Chicago and a team of collaborators to measure metabolic rates and body composition in zoo-living apes across the U.S. The results were eye-opening: even in captivity, gorillas and orangutans average only 14 to 23 percent body fat and chimpanzees less than 10 percent, on par with Olympic athletes.

Among our primate cousins, we humans are clearly the odd ape out. Somehow humans evolved to require much higher levels of physical activity for our bodies to function normally. Sitting for hours on end, grooming and napping (or watching the tube) have gone from standard practice to a health risk. When did we trade the low-key existence of our fellow apes for a more strenuous way of life and why? Fossil discoveries are helping to piece the story together.

BRANCHING OUT

OUR LIMB OF THE PRIMATE FAMILY TREE, the hominins, split from that of chimpanzees and bonobos about six million or seven million years ago, near the end of Miocene geologic time period. Until fairly recently, there were few hominin fossils recovered from the earliest portion of the lineage. Then, in quick succession during the 2000s, paleoanthropologists working in Chad, Kenya and Ethiopia reported finds of three hominins from this critical period: *Sahelanthropus, Orrorin* and *Ardipithecus.*

Each of these early hominins is distinct from any of the living apes in the anatomical details of their cranium, teeth and skeleton. Nevertheless, aside from walk-

© 2019 Scientific American

ing on two legs, it appears these species lived a very apelike existence. Their molars were similar in size and sharpness to chimpanzees, with somewhat thicker enamel, suggesting a mixed diet of fruit and other plant foods. *Ardipithecus,* found in 4.4-million-year-old deposits in Ethiopia and by far the best known early hominin, had long arms, long, curved fingers and grasping feet, indicative of a life spent partly in the trees. New biomechanical analyses, led by my City University of New York graduate student Elaine Kozma, show that *Ardipithecus* had evolved changes in its pelvic anatomy to permit fully upright, energetically efficient walking without compromising the ability to power itself into the canopy. Our early ancestors were clearly comfortable in two worlds, on the ground and in the trees.

BINGEING ALL 63½ HOURS OF *GAME OF THRONES* WILL COST YOU ONE DAY ON THIS PLANET.

From about four million to two million years ago the hominin record is dominated by the genus *Australopithecus,* with at least five species recognized today, including the famous "Lucy" and her kin. Anatomical changes in the lower limb point to improved walking ability and more time on the ground compared with earlier species. The grasping foot is gone in *Australopithecus,* the big toe in line with the others, and the legs are longer, the same ratio of leg length to body mass that we see in living humans. Analyses of the pelvis by Kozma, together with recent work on the fossilized footprints from Laetoli in Tanzania, indicate that this creature had an effectively modern gait. Long arms and fingers tell us these hominins were still regularly in the trees to forage and perhaps to sleep. Analyses of the wear patterns on their teeth suggest *Australopithecus* species primarily ate plant foods, just as the earliest hominins did before them and living apes do today. Based on their large, thick-enameled molars, *Australopithecus* diets most likely leaned more on harder and more fibrous foods, particularly when preferred foods were not available.

The evolution of an upright, striding bipedal gait in these early hominins is important, indicating a different approach for navigating their landscape. Covering more ground for fewer calories might have enabled these species to expand their range and thrive in less productive habitats than apes today. There are other notable and intriguing changes, too, such as the loss of big, sharp canines in males, which seem to reflect changes in social behavior. Yet the plant-based diet and retained climbing adaptations tell us their foraging ecology and daily activity remained quite apelike. Distances traveled per day were probably modest, with lots of time spent resting and digesting bellyfuls of fibrous plant food. It is unlikely they needed, or often got, their 10,000 steps a day.

Some two million years ago the telltale signs of curious or clever hominins experimenting with new ideas and approaches began to emerge. In 2015 Sonia Harmand of Stony Brook University and her team recovered large, unwieldy stone tools, some weighing more than 30 pounds, from 3.3-million-year-old sediments on the western shore of Lake Turkana in Kenya. In the past 15 years excavations at 2.6-million-year-old sites in both Ethiopia and Kenya have found stone tools associated with fossilized animal bones bearing the unmistakable gouges and scrapes of butchery. By 1.8 million years ago cut-marked bones and stone tools were the norm, and it was not just the sick and injured animals that fell prey to these hominins. Analyses of butchered bones at Olduvai Gorge in Tanzania show that prime-aged ungulates were targeted. Just as important, unlike every hominin before, by 1.8 million years ago hominins had expanded outside of Africa into Eurasia, from the foothills of the Caucasus Mountains to the rain forests of Indonesia. Our predecessors had jumped the ecological fence and were capable of thriving nearly anywhere.

Forget the tales of some clandestine meeting in the Garden of Eden or of Prometheus doling out fire. It was this million-year dalliance with stones and meat and the development of a hunting-and-gathering strategy that pushed our lineage away from the other apes, changing things irrevocably. This tectonic shift marked the evolutionary emergence of us, the genus *Homo*.

FOOD FOR THOUGHT

IN ECOLOGY AND EVOLUTION, diet is destiny. The foods animals eat do not just shape their teeth and guts but their entire physiology and way of living. Species evolved to eat foods that are abundant and stationary need not roam too far or be too clever to fill up; grass does not hide or run away. Eating foods that are hard to find or capture means more travel, often coupled with increased cognitive sophistication. For instance, fruit-eating spider monkeys in Central and South America have larger brains and travel five times farther every day than the leaf-obsessed howler monkeys that share their forests. Carnivores on the African savanna travel three times farther a day than the herbivores they hunt.

Shifting from the pure gathering lifestyle of apes and early hominins to the hunting-and-gathering strategy that marks the genus *Homo* had major ramifications. It made these social primates even more tightly knit. Relying on meat requires cooperation and sharing—and not just because you cannot kill or eat a zebra by yourself. Meat is difficult to obtain, and sharing more predictable plant foods is what allows hunting and gathering to work. Hunter-gatherer popula-

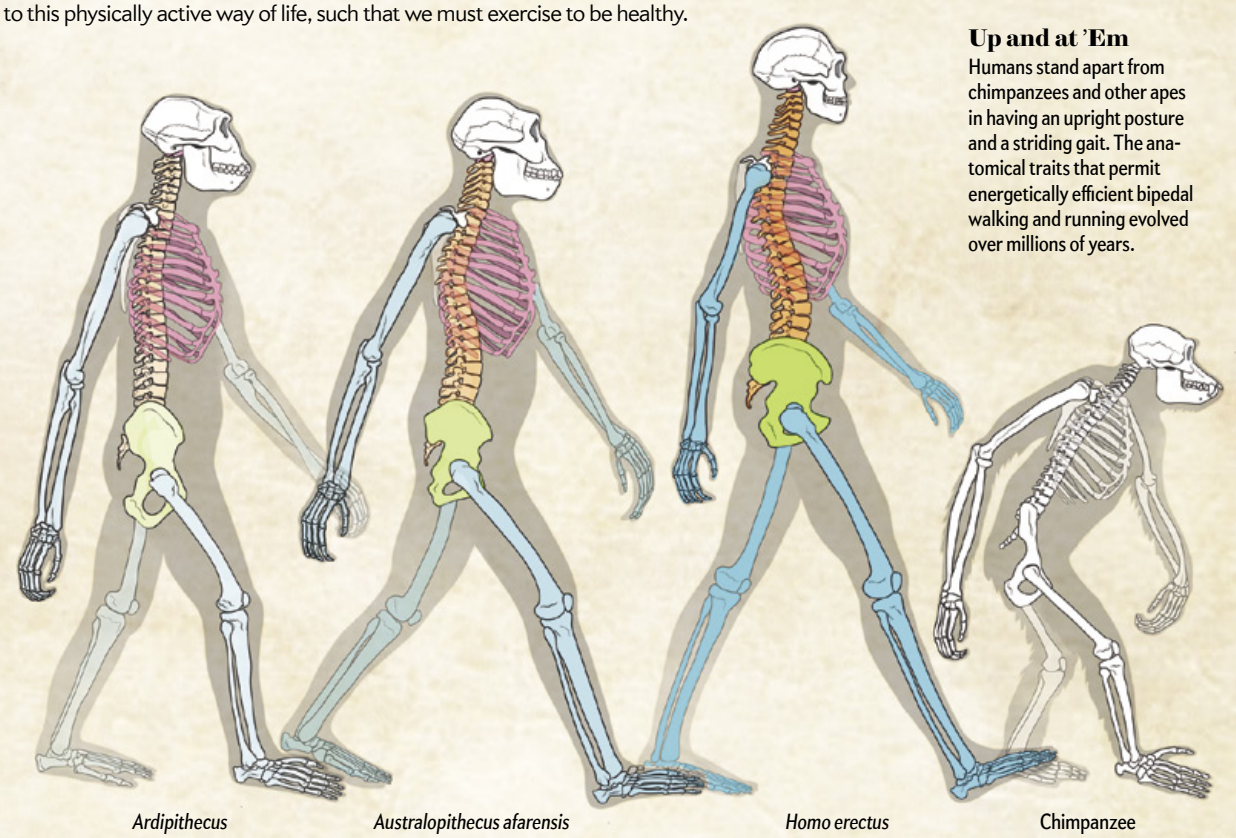
© 2019 Scientific American

Built to Move

As hominins evolved anatomical changes that facilitated upright walking (*shown*), they were able to cover more ground for fewer calories, allowing them to expand into new habitats. The subsequent advent of hunting further increased activity levels of hominins, requiring them to travel farther to find food. Our physiology has adapted to this physically active way of life, such that we must exercise to be healthy.

Up and at 'Em

Humans stand apart from chimpanzees and other apes in having an upright posture and a striding gait. The anatomical traits that permit energetically efficient bipedal walking and running evolved over millions of years.



Ardipithecus | *Australopithecus afarensis* | *Homo erectus* | Chimpanzee

Piecemeal Evolution

Early human ancestors, including *Ardipithecus* and *Australopithecus*, walked on two legs yet also retained adaptations to climbing in the trees, ate a plant-based diet like today's apes do and lived exclusively in Africa. By 1.8 million years ago *Homo* had evolved modern body proportions, adopted a dietary strategy of hunting and gathering, and spread out of Africa into Eurasia.

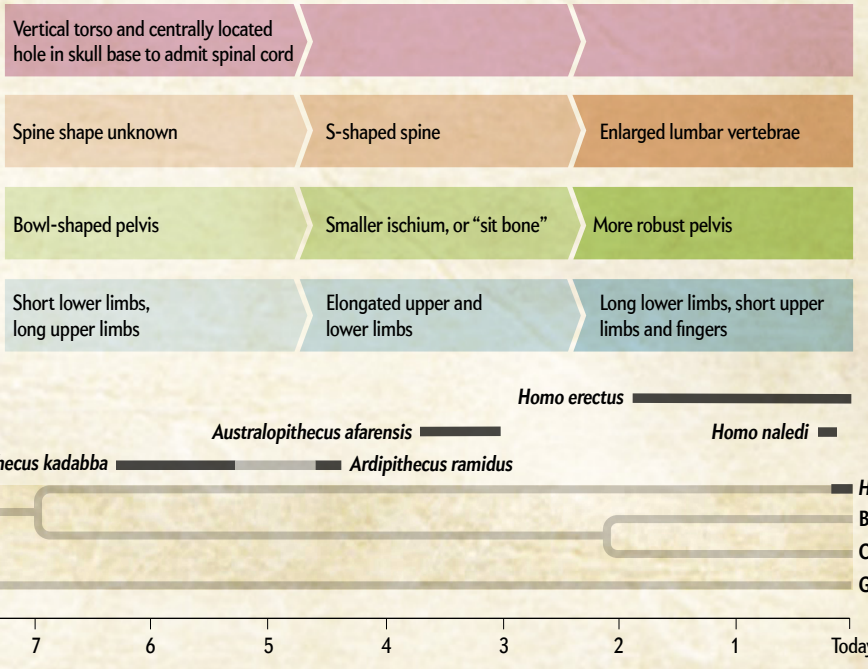


Illustration by Mesa Schumacher

© 2019 Scientific American

tions today get roughly half their daily calories from plants. Recent analyses of the food trapped inside their fossilized dental calculus show that even Neandertals, masterful hunters and avatars of vegetable-boycotting Paleo diet bros everywhere, ate a balanced diet, with plenty of plants, including grains.

Hunting and gathering also put an evolutionary premium on intelligence. Technological innovation and creativity meant more calories and better odds of reproduction. Social intelligence would have proved invaluable, as coordination and communication became ingrained parts of the hominin strategy. Discoveries by Alison Brooks of George Washington University, Rick Potts of the Smithsonian National Museum of Natural History and their colleagues at the site of Olorgesailie Basin in Kenya, published in 2018, show that by 320,000 years ago, hominin cognition had blossomed into the kind of sophistication seen in modern humans, with black and red pigments for visual expression and long-distance trade networks for premium stone tool material. The age of these finds corresponds well with the oldest *Homo sapiens* fossils found to date, reported in 2017 from the 300,000-year-old site of Jebel Irhoud in Morocco.

EXERCISE IS NOT OPTIONAL; IT IS ESSENTIAL.

Moreover, hunting and gathering required hominins to work harder for their food. Simply moving up the food chain means food is harder to find; there are a lot more plant calories on the landscape than animal calories. Hunter-gatherers are remarkably active, typically covering nine to 14 kilometers a day on foot—about 12,000 to 18,000 steps. Work that David Raichlen of the University of Arizona, Brian Wood, now at the University of California, Los Angeles, and I have done with the Hadza hunter-gatherer population in northern Tanzania shows that men and women in that group log more physical activity in a day than Americans typically get in a week and travel three to five times farther every day than any of the great apes. Early members of our genus, without the benefit of technological innovations such as the bow and arrow, might have been even more active. In a landmark paper in 2004, Dennis Bramble of the University of Utah and Daniel Lieberman of Harvard argued that our genus evolved to run prey to exhaustion, pointing to a number of features in the *Homo erectus* skeleton that appear to reflect endurance running.

The steady increases in brain size and technological complexity over the past two million years seem to accumulate like a snowball rolling downhill, but any impression of momentum is an illusion. Evolution has a great memory but no plans. In 2015 Lee Berger of the University of the Witwatersrand in South Africa and his team announced their discovery of hundreds of fossils of *Homo naledi,* a new species recovered from deposits deep in the Rising Star Cave system in South Africa, dated to between 236,000 and 335,000 years old. With a brain size only 10 percent larger than *Australopithecus* and a body size similar to early *Homo,* this hominin appears to represent a lineage within our genus that stalled out in the early Pleistocene, persisting quite happily for more than a million years without the continued increase in brain size seen in other *Homo* species. *H. naledi* is an important reminder that evolution is not trying to get anywhere. We were not inevitable.

SHARKS ON THE SAVANNA

NO TRAIT EVOLVES IN ISOLATION: brains must fit snugly inside their skulls, teeth inside their jaws; muscles, nerves and bones must function harmoniously. Behavioral traits are no different. When a behavioral strategy—such as hunting and gathering—becomes the norm, physiology adapts to accommodate and even depend on it.

Take vitamin C, for instance. Early mammals evolved a multistep process to make this crucial nutrient on their own, a cascade involving several genes that remains functional in rodents, carnivores and many other mammals. Tens of millions of years ago our primate ancestors became so fixated on eating fruits rich in vitamin C that making their own became an unnecessary cost. Their physiology adapted to their behavior, with mutations accumulating in the gene needed in the final step of synthesis. Consequently, today's anthropoid primates—monkeys, apes and humans—cannot make vitamin C. Without it in our diets, we get scurvy and die.

Further afield, yet closer to home, is the evolution of a specialized form of breathing called ram ventilation in several species of sharks and scombrid fish (the group that includes tuna and mackerel). These lineages evolved highly active foraging behavior, swimming nonstop day and night. Their anatomy and physiology adapted, using the constant forward motion to ram water into their mouths and past their gills. This change eliminated the need to pump water past the gills, leading to the evolutionary loss of the associated gill musculature. This loss saved energy but left these species vulnerable to suffocation. If they stop moving, they die.

Although we have long known that exercise is good for us humans, we are only beginning to appreciate the myriad ways our physiology has adapted to the physically active way of life that hunting and gathering demands. Nearly every organ system is implicated, down to the cellular level. Some of the most exciting work in this area has focused on the brain. For one thing, our brain has evolved to get less sleep, even in societies without artificial lighting or other modern nighttime distractions. Humans around the globe—whether it is the Hadza on the African savanna, the Tsimane horticulturalists in the Amazonian rain forest or urbanites in New York—clock about seven hours of sleep a night, far less than apes.

© 2019 Scientific American

Raichlen and his colleagues have shown that our brain has evolved to reward prolonged physical activity, producing endocannabinoids—the so-called runner's high—in response to aerobic exercise such as jogging. Raichlen and others have even argued that exercise helped to enable the massive expansion of the human brain and that we have evolved to require physical activity for normal brain development. Exercise causes the release of neurotrophic molecules that promote neurogenesis and brain growth, and it is known to improve memory and stave off age-related cognitive decline.

Our metabolic engines have evolved to accommodate increased activity as well. Humans' maximum sustained power output, our VO_{2max}, is at least four times greater than that of chimpanzees. This increase stems largely from changes in our leg muscles, which are 50 percent bigger and have a much greater proportion of "slow-twitch" fatigue-resistant fibers than the leg muscles of other apes. We also have more red blood cells to carry oxygen to working muscles. But the adaptations to exercise appear to go even deeper, accelerating the rate with which our cells function and burn calories. My work with Ross, Raichlen and others has shown that humans have evolved a faster metabolism, providing fuel for increased physical activity and the other energetically costly traits that set humans apart, including bigger brains.

All of this evidence points toward a new way of thinking about physical activity. Since the sweaty spandex excitement of the 1980s, exercise has been sold as a way to lose weight or as a health-conscious buffet item to add to our lifestyle, like oat bran muffins. But exercise is not optional; it is essential, and weight loss is probably the one health benefit it largely fails to deliver. Our bodies are evolved to require daily physical activity, and consequently exercise does not make our bodies work *more* so much as it makes them work *better*. Research from my lab and others has shown that physical activity has little effect on daily energy expenditure (Hadza hunter-gatherers burn the same number of calories every day as sedentary Westerners), which is one reason exercise is a poor tool for weight loss. Instead exercise regulates the way the body spends energy and coordinates vital tasks.

Recent advances in metabolomics have shown that exercising muscles release hundreds of signaling molecules into the body, and we are only beginning to learn the full extent of their physiological reach. Endurance exercise reduces chronic inflammation, a serious risk factor for cardiovascular disease. It lowers resting levels of the steroid hormones testosterone, estrogen and progesterone, which helps to account for the reduced rate of reproductive cancers among adults who exercise regularly. Exercise may blunt the morning rise in cortisol, the stress hormone. It is known to reduce insulin insensitivity, the immediate mechanism behind type 2 diabetes, and helps to shuttle glucose into muscle glycogen stores instead of fat. Regular exercise improves the effectiveness of our immune system to stave off infection, especially as we age. Even light activity, such as standing instead of sitting, causes muscles to produce enzymes that help to clear fat from circulating blood.

No wonder populations such as the Hadza do not develop heart disease, diabetes or the other maladies that afflict industrial countries. But we do not need to cosplay as hunter-gatherers or run marathons to reap the benefits of a more evolutionarily informed life. The lesson from groups such as the Hadza, Tsimane and others is that volume matters more than intensity. They are on their feet and moving from sunrise until dusk, racking up more than two hours of physical activity a day, most of it as walking. We can emulate these same habits by walking or biking instead of driving, taking the stairs, and finding ways to work and play that keep us off our butts. A recent study of Glaswegian postal workers shows us what this can look like. These men and women were not committed athletes but were active throughout the day, handling the mail. Those who got 15,000 steps or spent seven hours a day on their feet (numbers similar to what we see with the Hadza) had the best cardiovascular health and no metabolic disease.

While we are at it, we might take other lessons for living well from groups like the Hadza. Beyond the copious amounts of exercise and whole food diets, daily life for these cultures is full of fresh air, friendships and family. Egalitarianism is the rule, and economic inequality is low. We do not know exactly how these factors affect the health of hunter-gatherers, but we know their absence contributes to chronic stress in the developed world, which can promote obesity and disease.

Embracing more physically active life habits would be easier if we did not have to wrestle with the 400-pound gorilla in our head. Like vitamin C for our anthropoid ancestors, exercise was unavoidable and plentiful during the past two million years of hominin evolution. There was no need to seek it out, no evolutionary pressure to lose the ancient, simian weakness for gluttony and sloth. Today, as masters of our environments, we are giving our inner apes too much say in how the modern world is engineered: filling up on easy food, bingeing *The Walking Dead* instead of actually walking, sitting for hours at our desks grooming one another on social media. We are fascinated when we see ourselves in great apes, but we should worry when we see them in us. Underneath the surface, we are more different than we seem. SA

Herman Pontzer is an associate professor of evolutionary anthropology at Duke University. He studies how evolution has shaped human physiology and health.

MORE TO EXPLORE

The Crown Joules: Energetics, Ecology, and Evolution in Humans and Other Primates. Herman Pontzer in *Evolutionary Anthropology,* Vol. 26, No. 1, pages 12–24; January/February 2017.

Economy and Endurance in Human Evolution. Herman Pontzer in *Current Biology,* Vol. 27, No. 12, pages R613–R621; June 19, 2017.

scientificamerican.com/magazine/sa

© 2019 Scientific American

| GHOST | FLOWERS |

The genes of Hawaiian plants, extinct for more than a century, have been brought back from the dead. Today we can smell their scents

By Rowan Jacobsen

Photographs by Floto + Warner

© 2019 Scientific American

BACK FROM THE BRINK: The Wynberg conebush (*left*) went extinct in 1806, and Maui's mountain hibiscus (*right*) followed in 1912. But their DNA has been recovered, and some rejuvenated scent genes are once again producing fragrances. The hibiscus, sniffed by people for the first time in more than a century, evokes bark and juniper, with hints of citrus and thyme.

© 2019 Scientific American

IN 1912, ON THE ANCIENT LAVA FIELDS OF HALEAKALĀ ON THE HAWAIIAN ISLAND of Maui, a single tree stood near death. Fifteen feet tall, its bark encrusted with lichens, it was down to its last flower.

The Hawaiians called this tree *hau kuahiwi,* the mountain hibiscus. Unlike the more familiar Hawaiian hibiscus, which grows in moist valleys and opens wide in a welcoming *aloha,* the mountain hibiscus grew only on the dry, well-drained lava fields of Hawaii's volcanoes. The plant unfolded only two of its five hibiscuslike petals, keeping the rest closed in a demure, curved tube designed for Maui's honeycreepers—nectar-eating songbirds with curved bills that were its favored pollinators.

IN BRIEF

Dead is dead, went the dogma of extinction, for when the last of a species vanishes, it is gone from the world forever. **But genes of the dead** can be resurrected, a group of scientists has shown, by recovering DNA and making it function again. **Scent genes** from long-gone flowers have been recovered in this way, and researchers have smelled their products.

But this tree had not reproduced in years. Most honeycreepers had disappeared as the 19th century gave way to the 20th. The lava fields of Haleakalā had been turned into cattle ranches. Cows rubbed its trunk raw. Rats ate its seeds.

Up the slope came a botanist, dressed in Rough Rider cavalry hat and khakis, a collector's bag over his shoulder. His name was Gerrit Wilder, and he was on the original expedition that identified this tree in 1910. Because of that, the tree was named for him, *Hibiscadelphus wilderianus*. It was the only member of its species ever found. Its sickly state was why Wilder had returned. He plucked the last flower, along with some twigs and leaves, and nestled them into his bag. Then he turned and made his way slowly down the slope.

Not long after that, the tree succumbed to the cattle and the rats and dropped its final leaves. *H. wilderianus* was extinct. And that should have been that. Extinction is supposed to be forever.

Recent breakthroughs in DNA sequencing, however, have made it increasingly easy to read the genes of long-dead organisms and "reboot" those DNA stretches. Serious efforts are underway to use such tech to revive the passenger pigeon and the woolly mammoth. Both projects depend on bioengineering advances that are still years away. Yet in 2018, in an eighth-floor laboratory built above the burgeoning Seaport District in Boston, a crucial part of this long-dead mountain hibiscus came back to life.

A collection of gene engineers, working for a company called Ginkgo Bioworks, was able to re-create the scent genes from the flower. They rebuilt the genetic material that had produced the flowers' distinctive odor, got it working again in another life-form—a yeast—and human noses smelled something that had vanished from the planet more than a century ago. Like Odysseus raising the dead in Hades and plying them for information, some kind of communication took place between the living and the deceased. There were no flowers, no petals, but these were the actual DNA sequences of the plant telling cells to churn out molecules as they used to do on Maui, and those molecules were grabbed in people's noses, sending signals to their brains. It was the most tangible sign yet that the hard membrane of extinction is beginning to soften. This newfound porosity forces a strange question: Can we reboot enough genes to say that something isn't quite dead anymore?

SCENT OF LIFE

THE RESURRECTION PROJECT BEGAN, oddly enough, at the 2014 annual convention of the International Federation of Essential Oils and Aroma Trades in Rome, where Jason Kelly, Ginkgo's CEO, was looking to drum up business. Kelly and his Ginkgo co-founders graduated in 2008 from the Massachusetts Institute of Technology with some of the first Ph.D.s in synthetic biology awarded for that specialization. His firm is also highly specialized: if another company needs a new microbe to produce some valuable molecule—for fuel, fiber, fragrance, pharmaceutical, whatever—Ginkgo will design and test hundreds of prototypes in its biofoundries and hand over the top performers.

RECOVERED GENES: Researchers found fragmented DNA in extinct plant specimens at the Harvard Herbaria. They spliced the DNA together into scent genes for the Wynberg conebush (*Leucadendron grandiflorum* Salisb.) (*1*); the mountain hibiscus (*Hibiscadelphus wilderianus* Rock) (*2*); and the Falls-of-the-Ohio scurfpea (*Orbexilum stipulatum* [Torr. & A. Gray] Rydb.) (*3*). These three produced fragrant compounds. A myrtle, *Myrcia skeldingii* Proctor (*4*), yielded a gene that did not make a scent molecule.

© 2019 Scientific American

CULTIVATED
HERBARIUM OF THE ARNOLD ARBORETUM HARVARD UNIVERSITY
CULTIVATED PLANTS
UNION OF SOUTH AFRICA
KIRSTENBOSCH BOTANIC GARDENS
NEWLANDS, CAPE PENINSULA
Leucospermum grandiflorum (Salisb.) R. Dr.
Displayed at the 17th Int'l. Hort. Congress
at College Park, Maryland, August 15-20, 1964.
Comm. H.B. Sprecher
wm. F.D. Mayer & T.H. Dudley

1

2

HARVARD
Orbexilum stipulatum (T. & G.) Rydb.
Grimes '58
Orbexilum stipulatum (T. & G.) Rydb.
J. & C. Henkle, 11 Jan. 1906
Dot. Kevin L. Thomas, 28 Feb. 1983
Ky. Botanical Gardens, Louisville

3

HARVARD
INSTITUTE OF JAMAICA
FLORA OF JAMAICA
Name: Myrcia sheldonii Proctor
Locality: Mason River district, c. 3 miles due northwest of KELLITS.
Altitude: 2100-2300 ft. Date: July 4, 1948
Habitat: Thickets along bank of stream.
Collector: George R. Proctor

4

SPECIMENS PHOTOGRAPHED AT HERBARIUM OF ARNOLD ARBORETUM AND GRAY HERBARIUM OF HARVARD UNIVERSITY

© 2019 Scientific American

Many of Ginkgo's best clients are in the flavors and fragrances industry, where raw ingredients can be astronomically expensive. All those fragrance molecules are produced by enzymes in the plant cells, and the blueprint for those enzymes is coded in DNA by a gene. Like software, this code can run on any compatible platform, and life is surprisingly platform-agnostic. All living things use the same four-letter language of DNA—components labeled A, T, C and G—and yeast and plants run many of the same genes. By inserting fragrance genes into specially engineered strains of brewer's yeast, Ginkgo brews scent molecules in a flask, just like making beer.

At the trade show, Kelly met a consultant for Givaudan, the Swiss perfume giant, who told Kelly about Givaudan's Scent Trek program, which dispatched explorers into the world's rain forests to capture the air around rare flowers so the scents could be identified. Kelly was intrigued. If Ginkgo could get samples of these plants, the company could sequence the genes and synthesize the enzymes that made the smells. But as the two brainstormed, Kelly had a much crazier idea. What if he was able to go beyond obscure plants and bring back the scent of flowers that no longer even existed?

This would be the first step, he thought, in reversing a tremendous biological waste. "The planet has spent three billion years trying out different DNA sequences through this process we call evolution," Kelly says, "and that's what we have today. But along the way, a lot got lost for some random reason—a meteor or whatever—and some of that stuff was incredible. The planet spent hundreds of millions of years evolving DNA. And we just have to let it go away? For a biological designer, it's frustrating to imagine losing all that great code."

"The planet has spent hundreds of millions of years evolving DNA. And we just have to let it go away? For a biological designer, it's frustrating."
—Jason Kelly, Ginkgo Bioworks

Kelly's original plan riffed on *Jurassic Park:* Recover an Ice Age flower from the Arctic permafrost, sequence its genes and synthesize the ones responsible for fragrance, then put them in yeast cells. When the genes instructed the cells to make the fragrant molecules, Kelly could brew up a little Extinction N° 5.

It was a long shot. Although a handful of ancient genes had been reconstructed in labs, most simply sat there, never being asked to produce a protein and thus rejoin our world. Even if Ginkgo was able to rebuild old genes, those genes might not function in new yeast. Kelly also worried about tying up precious resources. Everyone at his company was already overworked. The last thing they needed was to get sucked into a Jurassic lark.

But the project found a champion in Christina Agapakis, Ginkgo's creative director. Agapakis earned her Ph.D. in synthetic biology at Harvard University, and she worked on optimizing bacteria to produce hydrogen fuel and created art based on the shapes of antibodies. Warm and witty, she was drawn to research that probed the borderlands between natural and unnatural and opened up interesting conversations about genetically modified organisms. A perfume of extinct florals that people could smell while meditating on these lost species was right up her alley. She dubbed the venture "Project Cretaceous," after the period when flowers first came into existence. She began by contacting experts in Ice Age excavations, who told her it was impossible to sequence a full genome out of the gunky specks of plant that emerge from the permafrost. The Ice Age was a dead end.

Before giving up, Agapakis did what any good Millennial would do: she googled "extinct plant DNA sequencing." Far down the list of results, she found an obscure paper from the *Biological Journal of the Linnean Society* on museomics, a new technique for extracting DNA from museum-preserved plants and animals. So she did not need permafrost after all. She just needed an herbarium.

The realization made the Harvard grad smile. She knew just where to find one of those.

THE DNA SEARCH

THE HARVARD HERBARIA, which date back to 1842, anchor one brick-lined end of a street named Divinity Avenue, and their numerous floors are filled with formaldehyde-scented cabinets holding more than five million samples. They do not embrace change enthusiastically, so when Agapakis pitched her plan in 2016, the curator was skeptical. Do *what* with their plants? The herbaria were not in the business of giving away their collection to for-profit entities. Besides, they had no searchable database for their holdings, so they had no idea if they had any extinct plants or not.

It took Agapakis months of negotiations to reach an agreement. The deal was sealed when she offered to provide genomes of any extinct plants she found to the research community. Even so, she had to find the plants on her own, with no help from herbaria staff, and if she did find what she was looking for she could not take more than a pinky-nail-sized fragment of extraneous material.

Agapakis and Dawn Thompson, Ginkgo's head of Next Generation Sequencing, printed out the IUCN Red List of 116 modern plant extinctions and began their quest. The collection was arranged by plant family first and geography second, so the only way to find a sample was to go to the corresponding floor of the herbaria, find the aisle for the right family and then search in all the folders for the particular country or area. The aisles were endless, the cabinets seemingly filled with everything but the plants they were looking for. Then, in the Hawaii room, Agapakis cranked a big wheel to roll the creaking cabinets apart, opened the doors, paged through the folders, opened one, and gazed down at three long twigs holding an array of broad, beautiful leaves and a single pressed flower bud. "Flora of the Hawaiian Islands" read the attached card. "*Hibiscadelphus wilderianus.*" Agapakis felt an electric thrill. It was Wilder's extinct tree, right in front of her.

In the end, the scientists found 20 of the plants on their list in the herbaria, 14 of which had enough material to spare. Under the baleful eye of the curator, they snapped off the least important bits and placed them in plastic baggies.

© 2019 Scientific American

Road to Resurrection

Bringing genes back from the dead is not simply a matter of finding ancient DNA. The genetic material has to be refashioned into a working gene that can instruct a cell to make molecules. The gene—in this case from an extinct Hawaiian plant called a mountain hibiscus—then has to be placed into a new, living host. Synthetic biologists at Ginkgo Bioworks inserted the gene into yeast, where it actually made fragrant compounds.

1. From dried plant samples, researchers extract tiny DNA fragments.

2. The fragments are analyzed in a sequencer machine, which reads the order of their component nucleotides (dubbed A, T, C and G).

3. The gene that researchers want to re-create is for sesquiterpene synthase (SQS), the enzyme that assembles most floral scent molecules. Scientists take the ancient sequences and find some that match parts of an SQS sequence based on a gene in a current organism. The living sequence becomes a template to determine the position of each fragment.

4. The reconstructed sequence is turned into real DNA with a DNA printer, which builds out the molecule one component at a time. The end product is an SQS gene.

5. The gene is inserted into a yeast "host" engineered to accept it, and the yeast colonies grow in small wells. The instructions coded into the SQS gene tell the yeast to make a scent molecule, just as they once instructed the ancient hibiscus.

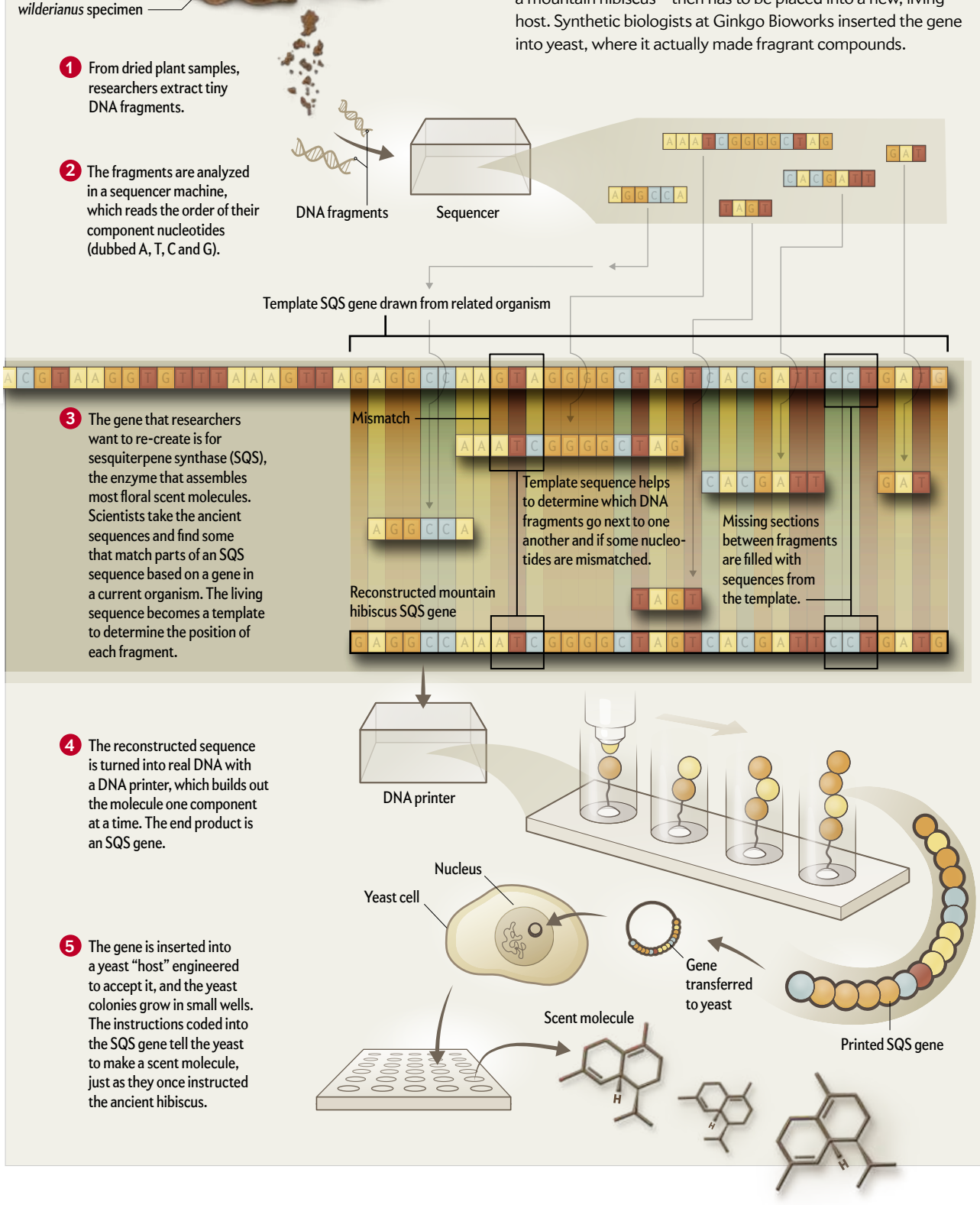


Illustration by Rebecca Konte

© 2019 Scientific American

1

2

3

DE-EXTINCTION ROOM: At Ginkgo Bioworks in Boston, a laboratory (*1*) is set up to engineer yeast to produce nonyeast molecules—such as scent compounds of extinct plants. An engineer there (*2*) prepares cells so their genetic code can be analyzed. Finally, Christina Agapakis, who directed the restoration project (*3*), sniffs mixtures of different compounds that originated from the recovered DNA of the extinct mountain hibiscus.

© 2019 Scientific American

Then it was time for the hard part. DNA degrades after an organism dies. Ginkgo was going to have to find needles of DNA in haystacks of cellulose. And the team had only enough material for a few attempts. The researchers decided to practice on an oak leaf scavenged from the streets of Boston. Even that did not go well. Despite their state-of-the-art sequencing equipment, they struggled to extract DNA from the samples. The ancient samples did not produce anything.

With pressure mounting to yield the sequencing machine to paying projects, Agapakis and Thompson had a sobering conversation. If they kept trying, they were going to run out of plant material, and there was no way they were getting more. They decided to put Project Cretaceous on hold until they found a more effective way of doing it.

Months later, at a conference, Kelly met Beth Shapiro, co-director of the Paleogenomics Lab at the University of California, Santa Cruz. This is the place you go if you want to de-extinct a mammoth or a passenger pigeon. Every year it gets better and better at extracting tiny amounts of DNA from iffy old material. In 2016 the lab was able to identify 0.01 to 0.05 percent mammoth DNA—a mere whiff of pachyderm—in 5,650-year-old lake sediments from an island in the Bering Strait. Send us your flowers, Shapiro said.

Thompson overnighted her plastic baggies of leaf matter to Josh Kapp, a grad student in the Paleogenomics Lab. Kapp did not like what he saw. He pulverized each sample into powder to maximize the surface area, but the plants did not powder as nicely as the bone he was used to. But after many filtration steps and some creative applications of chemicals that bind to DNA fragments, Kapp ended up with 14 microtubes holding the secrets of lost plants, which he packed in dry ice and sent back to Ginkgo. When Thompson ran the samples through her sequencing machine in Boston, she was thrilled to see numerous short reads come through: millions of fragments of genetic code, each just 40 or 50 letters long.

RECONSTRUCTION IN ACTION

BUT DID ANY OF THOSE fragments belong to a scent gene, and could they be put back together? Ginkgo was looking for genes, typically about 1,700 letters long, that made enzymes called sesquiterpene synthases (SQSs); these are the enzymes that stitch together most good floral scent molecules. A typical flower might have several of these genes. With all the tiny fragments they had recovered, it was as if the Ginkgo researchers had a book for each plant—the extinct plant genome—all chopped up into random 50-letter chunks and mixed together, and they needed to reassemble a few 1,700-letter passages in just the right way.

If the scientists had copies of the original books to use as a template or even a few chapters, they could figure out where the fragments went. Here evolution came to the rescue. It never invents anything from scratch. New species evolve from older species, tweaking or repurposing the original genes. So most SQS genes in modern plants share a lot of DNA code with closely related ancestors. Jue Wang, a computational biologist then working at Ginkgo, was tasked with this book-reassembly problem. He realized those modern SQSs could serve as the template. It was like trying to reconstruct a lost version of the Bible using the King James and New International versions as guides. The wording would not exactly match, but they would be a decent guide to what went where.

Bit by bit, Wang built his genes on the scaffolding modern relatives provided, relying on sequence overlaps for placement. He filled in any missing letters of DNA from the modern templates. If his fragmented Bible read "In th_ beg_ _ning was the W_ _ d," he could look to the King James and be pretty confident which letters were missing.

Ultimately Wang was able to reconstruct 2,738 versions of genes from the extinct flowers. Undoubtedly these strings of biological letters had a few typos. Would that ruin their functions? Occasionally a single wrong letter of DNA will break a gene catastrophically, as in sickle cell anemia. But often minor changes do not affect the end product. In fact, sometimes genes with significantly different forms will function similarly. In Biblical terms, "In the beginning was the Word" (King James) and "The Word was first" (The Message) do not match letter by letter, but both get the job done. Wang thought most of his letter strings would be too glitchy. He just hoped a few would work as instructions for a real cell.

For that to happen, these genes, which existed solely in Wang's computer, had to be converted into physical DNA. That is a fairly straightforward job, done with a DNA printer that resembles a 3-D printer but shoots out As, Cs, Gs and Ts, which bind together chemically into a classic double helix. Although this is often called synthetic DNA, it is just as real as any other DNA. Molecules are molecules.

Then it was up to the yeast, each of the 2,000-odd genes going into a colony bred to accept new DNA and make molecules according to its instructions. For several days the colonies frothed like beer brewing in their tiny containers. Scott Marr, a molecular microbiologist at Ginkgo, watched, wondering what they had made. When the fermentation subsided, Marr ran a sample of each colony through a mass spectrometer, a kind of artificial nose that was capable of detecting and identifying the minuscule amount of molecules being produced in each strain. Each mass shows up as a differently sized peak on a graph. It was Marr's job to read the pattern of peaks like a fingerprint.

He wrote programs to eliminate all the regular products of yeast metabolism in the machine's readout, so only nonyeast SQS products—scent-making sesquiterpenes—would show up. Mindful of the long odds and the likelihood of typos in the Ginkgo translations, Marr crossed his fingers and ran the samples. The readouts showed nothing. Then more nothing. It looked like the scientists had pieced together book passages with too many letter mistakes, paragraphs that no cell could read.

And then there it was: a peak. After a while, there was another and another. Marr let out a pent-up breath and began to match the molecular fingerprints to his database of terpenes. Then he broke the good news to the Project Cretaceous team: dozens of the flower-yeast chimeras were alive.

Agapakis sat at a table, listening to Marr's report and taking it all in. It had been three years since the initial crazy idea. Many times she and her colleagues had nearly abandoned it. And now they had molecules. Real molecules! Made by genes that had not existed in a century!

BACK IN THE REAL WORLD

GINKGO'S YEAST was able to get genes from three different extinct plants to produce sesquiterpenes. Although the microscopic amounts were far too small to smell directly, the scien-

© 2019 Scientific American

1

2

3

4

5

6

7

8

PROJECT CRETACEOUS: In the Ginkgo effort, which the researchers whimsically named for the Cretaceous period, they re-created genes for scent molecules from 14 extinct (or nearly so) plants. These eight yielded promising DNA sequences but did not churn out scent compounds when placed in designer yeast: *Erica pyramidalis Sol.* (*1*), *Crassula subulata var. subulata L.* (*2*), *Nesiota elliptica* (Roxb.) Hook. f. (*3*), *Pradosia glaziovii* (*Pierre*) T. D. *Penn.* (*4*), *Macrostylis villosa* (*Thunb.*) *Sond.* (*5*), *Shorea cuspidata P. S. Ashton* (*6*), *Stenocarpus dumbeensis Guillaumin* (*7*) and *Thamnea depressa* Oliv. (*8*).

© 2019 Scientific American

tists had some inklings of what the eventual floral nature might be, based on the smells of modern counterparts. One of the plants, the Falls-of-the-Ohio scurfpea—a legume that made the fatal mistake of growing only on a few rocky islets in the Ohio River that were drowned by dams in the 1920s—produced a handful of sesquiterpenes that, if some 21st-century relatives were any guide, would have woody, peppery, balsamic scents.

The Wynberg conebush—a five-foot-tall flower with white petals and a yellow head that grew in the granite hills above Cape Town until 1806, when it disappeared forever underneath South Africa's expanding vineyards—produced an astonishing 21 sesquiterpenes, many of which are associated with tantalizing scents: jasmine, lemongrass, cannabis, chamomile, turmeric, ginger, hops. That awkward mix sounded like a good match for a flower that had been noted for its "strong and disagreeable smell."

Eleven sesquiterpenes came from *H. wilderianus,* the mountain hibiscus, which had last released its essence to the world in 1912, as Gerrit Wilder picked that final flower and descended Haleakalā, never expecting that anyone would smell the *hau kuahiwi* again. From there, the genes' unlikely journey along Resurrection Road had taken them to the College of Hawaii's herbarium, where the plant was dried and pressed and eventually shared with the Harvard Herbaria. There it waited for decades for Agapakis to open its manila folder and break off a piece of the corpse. The genes were liquefied in Santa Cruz, digitized in Boston, then reanimated in the tender embrace of an organism completely unlike the one that hosted their last appearance on planet Earth. The genes had crossed time and space and outward form, but their information held.

Because of this work with ancient flowers, we are a tiny bit closer to coaxing saber-toothed tiger musk or Neandertal hemoglobin out of cells.

And then it was time to smell them. The Project Cretaceous team picked the *Hibiscadelphus* to go first because its allure was captivating, much as it had been for many a honeycreeper for millennia. On a bright August day in a crisp white conference room, the group gathered to sample a variety of formulations—created for the company by Berlin-based scent artist Sissel Tolaas—that blended the Hawaiian molecules in different combinations and concentrations. One of the molecules, juniper camphor, was a pricey ingredient in fragrant oils. *Hibiscadelphus* had expensive tastes.

They dipped paper fragrance test strips into 11 elfin bottles, held them a few inches from their noses and sniffed gently. Team members grinned at one another as if they could not quite believe they were here. "First resurrected fragrance!" Kelly announced. Agapakis's reaction was more visceral. "I feel overwhelmed," she said. "I couldn't imagine what this was going to smell like."

Some samples had flashes of citrus or thyme. All had a woody core of bark and juniper that must have been the essence of *hau kuahiwi*. "I like the lightness," Agapakis said, eyes closed as she inhaled. "It feels ethereal."

Lurking in the background of several samples was a smoky hint of sulfurous dirt. Kelly's eyes twinkled as he held one under his nose. "This is pretty magical, to be honest," he said. "I hope it captures people's imagination and gets them to think about what we've lost."

The scent—and the thoughts it inspires—is an important milestone, says Stanford University bioengineer Megan Palmer, a board member of Revive & Restore, a nonprofit that is supporting the passenger pigeon and woolly mammoth resurrection projects. "We can't know exactly what these flowers smelled like," she says, "but we can get molecular hints that we interpret through what we know about the species we see in the world today." As scientific advances, she adds, "these techniques can help us make smarter guesses at how extinct species functioned. They may even allow more ambitious projects to restore those functions and the species that gave rise to them."

Because of this work, we are a tiny bit closer to coaxing saber-toothed tiger musk or Neandertal hemoglobin out of cells. And as more of these freelance genes return to function in new forms, they make us begin to question our old emphasis on species. The traditional genetic container may not limit the life of its contents. For those sitting in that Boston conference room, it seemed clear that one of the most opportune moments in DNA's four-billion-year career had begun. This novel environment of bioengineering labs and digital databases and DNA printers was giving genes a newfound freedom to flow, new ways to replicate, new habitats to populate, new organisms to seduce. The original form may go extinct, but many functions can return, and at some point—nobody really knows what that is—that resurrection may get an organism to the point of "no longer dead."

As the essential oils saturated the air, the room became an unlikely tropical oasis, a hint of smoke in the distance, and it was easy to imagine the sun-baked lava fields of Haleakalā in the ancient past, a forest of mountain hibiscus all around, bright red honeycreepers flitting from blossom to blossom. That world will never come again, but some of the countless genes from primordial Hawaii and other lost landscapes may do just that. They are pressing against the membrane of extinction at this very moment, probing, hungry for any chance to get back in the action. SA

Journalist **Rowan Jacobsen** is author of several books, such as *Shadows on the Gulf* (Bloomsbury, 2011) and *The Essential Oyster* (Bloomsbury, 2016), and many magazine articles. He was a 2017–18 Knight Science Journalism Fellow at the Massachusetts Institute of Technology.

MORE TO EXPLORE

Timing and Causes of Mid-Holocene Mammoth Extinction on St. Paul Island, Alaska. Russell W. Graham et al. in *Proceedings of the National Academy of Sciences USA*, Vol. 113, No. 22, pages 9310–9314; August 16, 2016.

Natural Selection Shaped the Rise and Fall of Passenger Pigeon Genomic Diversity. Gemma G. R. Murray et al. in *Science*, Vol. 358, pages 951–954; November 17, 2017.

IUCN Red List of Threatened Species on *Hibiscadelphus wilderianus:* www.iucnredlist.org/species/30397/9536660

scientificamerican.com/magazine/sa

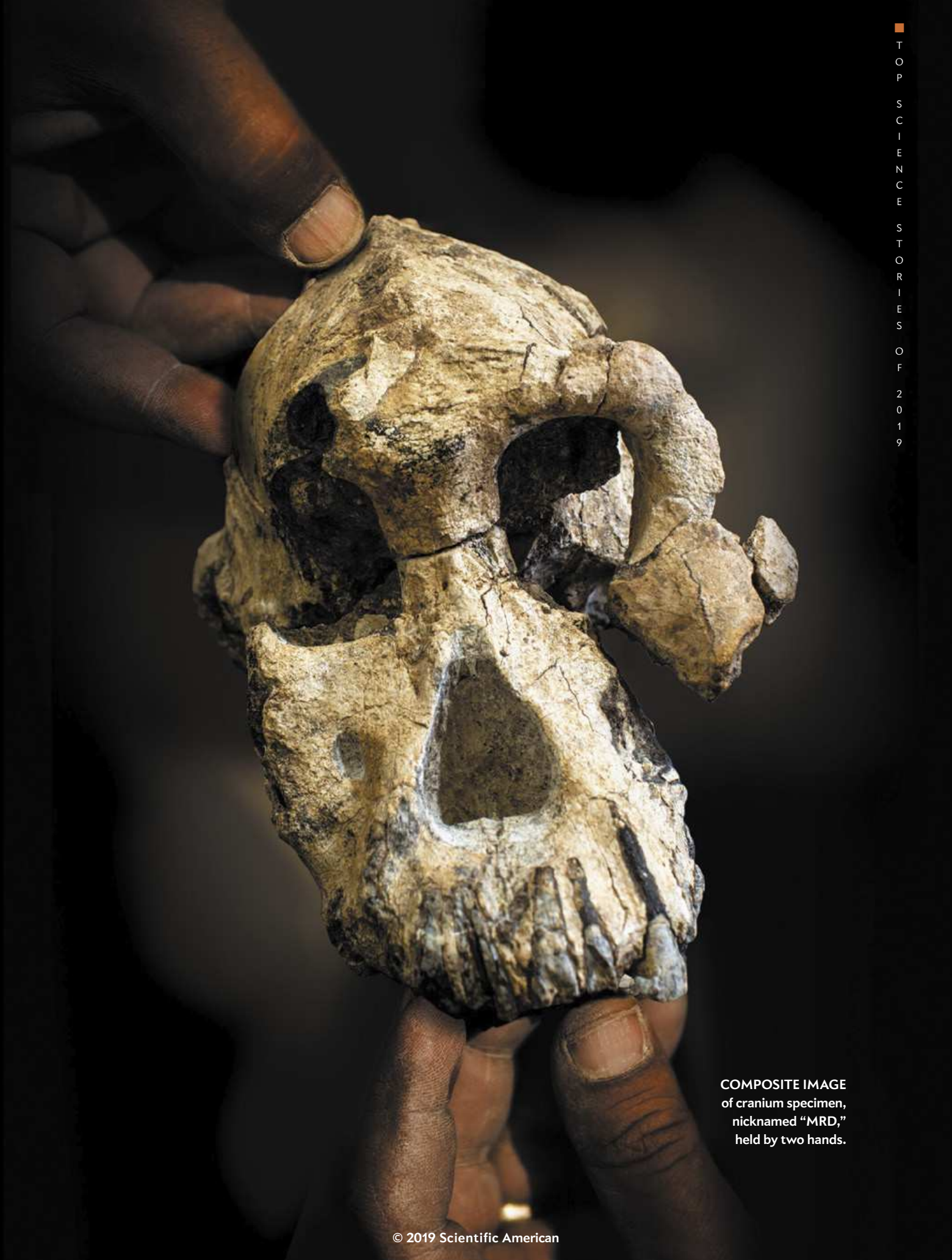
© 2019 Scientific American

A long-sought fossil cranium could redraw the human family tree

By Kate Wong

A FACE FROM DEEP TIME

© 2019 Scientific American



COMPOSITE IMAGE of cranium specimen, nicknamed "MRD," held by two hands.

© 2019 Scientific American

NEARLY 25 YEARS after researchers described the first fossil traces of *Australopithecus anamensis,* this unsung human ancestor is finally having its moment. Researchers working in Ethiopia have discovered a nearly complete fossil cranium of this long-vanished member of the hominin group that includes *Homo sapiens* and its close extinct relatives. The fossil, dated to 3.8 million years ago, reveals the never before seen face of *A. anamensis,* a species previously known mainly from jaws, teeth and a smattering of bones from below the head. Traits evident in the specimen hint that our family tree may need redrawing.

1

By some accounts, *A. anamensis* is the oldest unequivocal hominin, with some fossils dating from as far back as 4.2 million years ago. For years it has occupied a key position in the family tree as the lineal ancestor of *Australopithecus afarensis,* which is widely viewed as the ancestor of our own genus, *Homo.* Based on the ages and characteristics of the available fossils, scientists thought *A. anamensis* gave rise to *A. afarensis* through an evolutionary process termed anagenesis in which one species transforms into another. The new fossil throws a wrench into the works of that theory.

Yohannes Haile-Selassie of the Cleveland Museum of Natural History and his colleagues recovered the cranium from an area in northeastern Ethiopia's Afar region known as Woranso-Mille. Features of its teeth and jaws link it to the previously known fragmentary remains of *A. anamensis.* The fossil shows a creature with a projecting face, large canine teeth, flaring cheekbones, a crest atop its head that anchored strong jaw muscles, and a long, narrow braincase that held a brain the size of a chimpanzee's. The discovery team suspects the cranium belonged to an adult male *A. anamensis.*

Here is how it could upend the conventional wisdom: on the basis of the more complete *A. anamensis* anatomy seen in the newly discovered cranium, Haile-Selassie and his colleagues argue that an enigmatic 3.9-million-year-old forehead bone from the site of Belohdelie, also located in Ethiopia's Afar region, belongs to *A. afarensis.* If this supposition is right, *A. anamensis,* which is known from fossils spanning the time between 4.2 million and 3.8 million years ago, and *A. afarensis,* which apparently lived from 3.9 million to 3.0 million years ago, actually overlapped for at least 100,000 years in the Afar.

And that overlap would imply that *A. anamensis* could not have evolved into *A. afarensis* by means of anagenesis. Instead *A. afarensis* split off from *A. anamensis,* which continued to exist for a time alongside its daughter species. This branching mode of evolution, known as cladogenesis, can occur when populations of a species become isolated from one another and are thus able to evolve in different directions.

But the case for cladogenesis over anagenesis hinges entirely on that 3.9-million-year-old forehead bone from Belohdelie belonging to *A. afarensis*—no other *A. afarensis* remains recovered thus far are that old. And problematically, with only one *A. anamensis* forehead bone to compare it with—the one in the new fossil—one cannot exclude the possibility that other *A. anamensis* individuals might have had foreheads that looked like the Belohdelie one. Only the discovery of more hard-won fossil faces can resolve that unknown. SA

Kate Wong is a senior editor for evolution and ecology at *Scientific American.*

MORE TO EXPLORE

Tools from China Are Oldest Hint of Human Lineage Outside Africa. Colin Barras; ScientificAmerican.com, July 11, 2018.

Cave That Housed Neandertals and Denisovans Challenges View of Cultural Evolution. Kate Wong; January 30, 2019.

Is This the Footprint of One of the Last Neandertals? Kate Wong; February 15, 2019.

scientificamerican.com/magazine/sa

PRECEDING PAGES: DALE OMORI AND LIZ RUSSELL (*photograph*); JENNIFER TAYLOR (*composite image*); CLEVELAND MUSEUM OF NATURAL HISTORY; THIS PAGE: JOHN GURCHE (*facial reconstruction*); MATT CROW (*photograph*); CLEVELAND MUSEUM OF NATURAL HISTORY; MADE POSSIBLE THROUGH GENEROUS CONTRIBUTION BY SUSAN AND GEORGE KLEIN

© 2019 Scientific American

2

© YOHANNES HAILE-SELASSIE AND CLEVELAND MUSEUM OF NATURAL HISTORY

FACIAL RECONSTRUCTION of MRD specimen (*1*); MRD cranium specimen at Woransso-Mille site in Ethiopia (*2*).

© 2019 Scientific American

POLLUTION, like that in this neighborhood in Brooklyn, affects racial groups unequally.

JEFF SPIELMAN *Getty Images*

Air Inequality

U.S. racial minorities are exposed to more air pollution than white people yet cause less of it

By Andrea Thompson

HARLEM AND THE SOUTH Bronx have some of the highest asthma rates in New York City. And these predominantly black and Hispanic neighborhoods—studded with smokestacks and crisscrossed by gridlocked highways—are emblematic of a large body of research showing clear racial disparities in exposure to air pollution.

A study published in March in the *Proceedings of the National Academy of Sciences USA* shows that even though black and Hispanic people in the U.S. are exposed to more air pollution than white people, these groups consume less from the industries generating much of that pollution. The findings put hard data behind inequities that environmental justice advocates have reported on the ground, revealing that racial minorities bear a disproportionate amount of the costs of emissions tied to higher levels of consumption. "It echoes things we've been saying for decades in the environmental justice movement," says Kerene N. Tayloe, director of federal legislative affairs at nonprofit WE ACT for Environmental Justice.

The study also found that these disparities persist despite substantial overall reductions in air pollution in recent decades. Robert Bullard, a professor of urban planning and environmental policy at Texas Southern University, who was not involved with the research, says that this

© 2019 Scientific American

finding underscores the need to craft policies that reduce such pollution in the hardest-hit communities.

The researchers focused on fine particulate matter with a diameter of 2.5 microns or less ($PM_{2.5}$), generated by construction, fires and the combustion of fossil fuels. These particles can contain hundreds of different chemicals and can penetrate deep into the lungs, contributing to heart and lung disease. As part of its study, the team estimated that 102,000 people die prematurely every year from $PM_{2.5}$ emissions from human-made sources (as opposed to wildfires or other natural sources). That number is nearly double the amount of people who die annually from car crashes and murders combined, says study co-author Jason Hill, a biosystems engineer at the University of Minnesota.

To trace exposure to the root cause of emissions, the investigators mined publicly available economic and $PM_{2.5}$ data. They linked emissions from various sources (such as coal-fired power plants or diesel vehicles) to economic activity that largely drives these emissions (such as electricity use or food purchases). Next they apportioned this consumption among racial groups and found stark disparities: on average, black and Hispanic people are exposed to 56 and 63 percent more $PM_{2.5}$, respectively, than the amount generated by their consumption, whereas white individuals are exposed to 17 percent less (*graphic*). "It's a pretty dramatic difference," Hill says.

The analysis made clear that a wide range of economic activity contributes to $PM_{2.5}$ exposure. It also showed that the higher relative contribution from white people was not the result of buying more from highly polluting sectors but rather of higher overall consumption. Household income was a factor in the observed differences, but it did not account for the entire mismatch.

Even though the total number of deaths from $PM_{2.5}$ exposure has declined in recent years, most likely because of regulations, the idea "that everyone is receiving the same level of protection is blown away" by the new findings, Bullard says. The results from this study, he adds, emphasize the need to undo the legacy of previous policies and decisions that placed polluting infrastructure disproportionately in low-income and minority communities. "How can we somehow change this paradigm that it's okay to keep polluting communities of color and communities that have contributed least?" Bullard asks. Hill says authorities cannot simply target any one category of emitter, such as coal-powered plants or diesel vehicles; all the involved industries must be addressed.

Scientists, legislators and communities will need to jointly work out which policies and regulations can tackle overall pollution while reducing these inequities, Hill and others say. This new framework for tracking pollution exposure provides a tool to do that, says Anjum Hajat, an epidemiologist at the University of Washington School of Public Health, who was not involved with the study. It could, she says, be used to evaluate how regulations targeting certain air pollution sources actually reduce exposure for various groups—and could also potentially reveal any unintended consequences, such as shifting pollution to other areas.

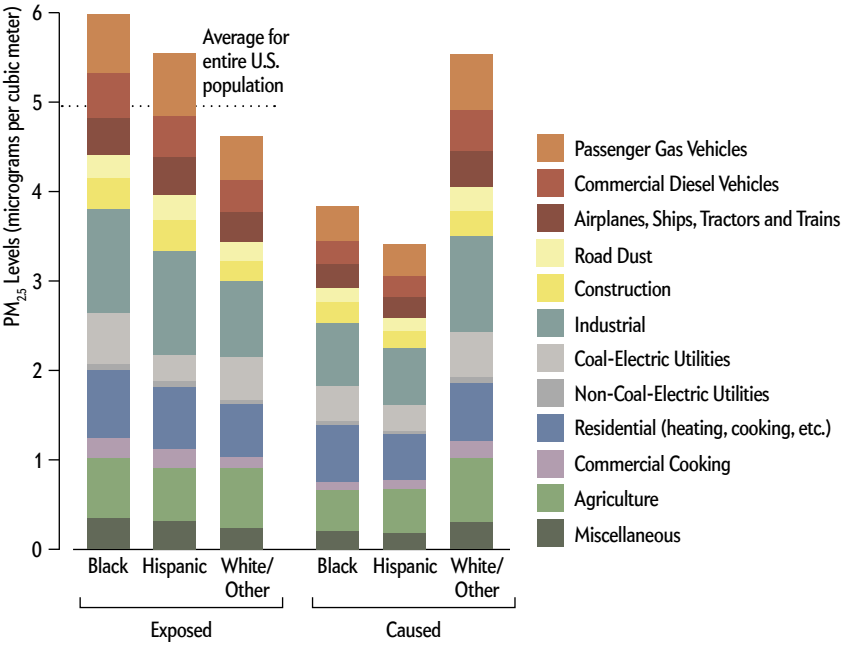
The work also underscores how important existing air pollution regulations are at a time when the Trump administration has moved to roll back several of them, including some governing pollution from coal-fired power plants and fuel efficiency in vehicles. When reached for comment, an Environmental Protection Agency spokesperson said the agency "will continue to monitor and report on key environmental indicators in low-income communities as a part of our commitment to healthy air, water and land for all Americans."

Research shows that even low levels of $PM_{2.5}$ can be harmful to the most vulnerable groups, such as children with asthma, so officials should be ratcheting up efforts instead of backing off, Hajat says. "We have seen this decline in [overall] air pollution over time" even as consumption has increased, she says, "so imagine what could happen if we really did make an effort as a society." SA

Andrea Thompson is an associate editor who covers sustainability at *Scientific American*.

Racial Groups' Exposure vs. Contribution to Air Pollution

Black and Hispanic individuals in the U.S. are exposed to higher levels of fine particulate matter ($PM_{2.5}$), on average, than white individuals yet consume less of the goods and services that cause such pollution. Black people, on average, experience the highest absolute pollution levels of the groups studied, whereas Hispanic people are exposed to the highest levels relative to their consumption.



SOURCE: "INEQUITY IN CONSUMPTION OF GOODS AND SERVICES ADDS TO RACIAL-ETHNIC DISPARITIES IN AIR POLLUTION EXPOSURE." BY CHRISTOPHER W. TESSUM ET AL., IN *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES USA*, VOL. 116, NO. 13; MARCH 26, 2019

Graphic by Melissa Thomas Baum, Buckyball Design

© 2019 Scientific American

THE LAST

Can we remove enough CO_2 from the atmosphere to slow or even reverse climate change?

By Richard Conniff

Photographs by Liz Tormes

INJECTION WELLS at Iceland's Hellisheiði geothermal power plant send brine down into deep bedrock, along with carbon dioxide pulled from the air.

© 2019 Scientific American

RESORT

© 2019 Scientific American

NOT LONG AGO IT SEEMED AS IF REDUCING GREENHOUSE gas emissions would be enough to save the world from climate change. Replace fossil-fuel power plants with clean energy sources, make cars and buildings more efficient, switch to LED lights, eat less meat, and so on. Slashing emissions and boosting renewables looked like the answer even to the Intergovernmental Panel on Climate Change as recently as 2005. But the strategy has not worked out as planned. Global emissions have gone up instead of down. It now appears that even cutting annual net emissions worldwide to zero by 2050 will not be enough.

To prevent economic and environmental devastation, climate experts maintain that we will now also have to achieve *negative* emissions. Doing so means removing billions of tons of carbon dioxide from the atmosphere every year. That is like saying we can no longer put out the garbage—and we need to steadily take back the garbage we put out in the past.

Negative emissions on a massive scale have become "a biophysical requirement" to meeting climate change goals, according to a 2018 study led by Jan C. Minx of Germany's Mercator Research Institute on Global Commons and Climate Change. Figuring out how to get there is a matter of "immediate urgency," he and his co-authors warn in *Environmental Research Letters,* if the world hopes to limit warming to 1.5 degrees Celsius. Almost every nation on the planet subscribed to that target—with a fallback of "well under" two degrees C—as part of the 2016 Paris Agreement on climate change. Warming is currently about one degree above preindustrial levels. But it is increasing at 0.2 degree C per decade. In a special report in October 2018, the Intergovernmental Panel on Climate Change warned that we have just 12 years to act if we hope to avoid slipping past 1.5 degrees C, the level regarded by most scientists as the furthest we can go if we hope to preserve life more or less as we know it.

Staying under that threshold mandates a specific "carbon budget," an overall amount of carbon dioxide we can add to the atmosphere without pushing warming beyond that temperature. At today's emissions—about 40 billion to 50 billion tons a year—"there may be only five years' worth of CO_2 emissions left" in the 1.5-degree scenario, Minx and his co-authors say. (For the remainder of this article, we use "tons" to mean "metric tons.") After that, every additional ton would require an equal withdrawal. His group estimates the world will need to remove 150 billion to more than one trillion tons of CO_2 from the atmosphere by 2100—roughly two billion to 16 billion tons a year, starting in 2050, with the number increasing significantly later in the century.

To do that, Minx and his team note, we will have to start building "several hundred" carbon capture and storage installations a year beginning in 2030, just 11 years from now. That could mean deploying big machines to pull carbon dioxide from the air or developing bioenergy power plants that burn trees—grown in continuous rotation—in a facility that captures emissions and sends them deep underground for permanent burial. Low-tech options would include replanting cut forests or expanding existing ones, improving farm and pasture soils so they hold more carbon, and crushing and spreading certain kinds of rock that soak up CO_2.

Most of the higher-tech carbon capture methods are still in the early stages of development, however. They require enormous investment at considerable risk of failure and entail major side effects, including competition for land that is already being used to feed people or provide habitat for wildlife.

And yet pursuing carbon capture on a massive scale appears to be our only option. When University of Washington statistician Adrian E. Raftery and his co-authors of a 2017 study in *Nature Climate Change* looked at current trends—not including negative emissions technologies—they found that we are on track to reach 3.2 degrees C of warming by the end of the century, with a range from two to 4.9 degrees C. In a subsequent study in the *Proceedings of the National Academy of Sciences USA,* Texas A&M climate scientist Yangyang Xu and a colleague catego-

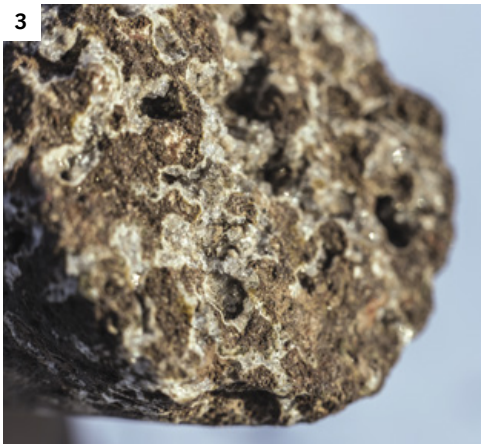
IN BRIEF

To limit global warming to 1.5 degrees Celsius, nations will have to remove one trillion tons of carbon dioxide emissions from the planet's atmosphere this century.

Finding the optimal mix of carbon capture methods will be critical. Machines that pull CO_2 from the air could remove 250 billion tons by 2100. Replanting clear-cut forests could achieve 180 billion tons.

Net costs range from $0 to $300 per ton. Unless big markets are developed to use the captured CO_2, a carbon tax may provide the best support for the techniques.

© 2019 Scientific American



CLIMEWORKS MACHINE extracts carbon dioxide from the atmosphere (*1*). An injector inside a nearby dome (*2*) sends the gas more than 700 meters underground, where it reacts with basalt bedrock to form carbonate, visible as white streaks in a core sample from the basalt (*3*).

rized warming greater than three degrees C as "catastrophic" and greater than five degrees C as posing "existential threats to a majority of the [human] population."

So let us assume that one trillion tons of negative emissions in this century—an average of 20 billion tons a year from 2050 to 2100—is necessary. What share of that pie could each method account for and at what cost? Given competition among methods for certain resources, such as land, what is the best mix to pursue? And can we muster the political will to pursue negative emissions—while drastically cutting our current carbon dioxide output?

CLEAN THE WIND

ON A HARDENED LAVA FIELD OF BOULDERS and moss in the foothills just outside Reykjavík, Iceland, a machine the size of a one-car garage pulls air through a chemical filter that extracts carbon dioxide. It is powered by waste heat from the geothermal power plant next door, and it pumps the captured carbon dioxide more than 700 meters underground, where the gas reacts with basalt rock and becomes solid mineral. Climeworks, a Swiss start-up, calls the operation the first direct air capture and storage plant in the world. It sequesters a modest 50 tons of carbon dioxide a year.

Direct air capture and storage may be the most straightforward path to negative emissions: banks of fans would harvest CO_2 from the sky and bury it. Scientific scenarios project that this technology could remove 10 billion to 15 billion tons of carbon dioxide a year by the end of the century; a few experts think 35 billion or 40 billion tons may be possible. This is such a tantalizing prospect that many climate scientists worry it could pose a moral hazard: people might think they can delay fossil-fuel reductions now in the hope of technological salvation later.

The most thorough review of removal methods—another 2018 study in *Environmental Research Letters*—takes a more sober view. Sabine Fuss of Mercator and her colleagues examined costs, side effects, environmental sustainability and other factors to project the carbon sequestration potential for seven major removal methods. Fuss and her co-authors put that potential for direct air capture in 2050 at only 500 million to five billion tons a year—adding up to 25 billion to 250 billion tons this century—at a cost of $100 to $300 per ton. For perspective, our cars each typically emit 4.6 tons of carbon dioxide a year.

Indeed, air capture "is not a silver bullet," says James S. Mulligan of the World Resources Institute's Food, Forests, and Water Program. "It's not a shiny object. It's kind of a cruddy object. But we need it." Some researchers claim they could get the cost below $100 per ton. Yet if the time line for doing that at scale is anything like the 60-plus years it took solar power to move from satellites in the 1950s to broad marketplace penetration today, Minx says "it may be too late."

Direct air capture also consumes enormous amounts of energy. Removing a million tons of carbon dioxide a year would require a 300- to 500-megawatt power plant, according to Jennifer Wilcox, a chemical engineering professor at Worcester Polytechnic Institute. If that were a coal-fired plant, it would create more emissions than it would remove. If power came from solar or wind farms, it would cover a lot of land that might already be in demand for farming or nature. And of course, a million tons

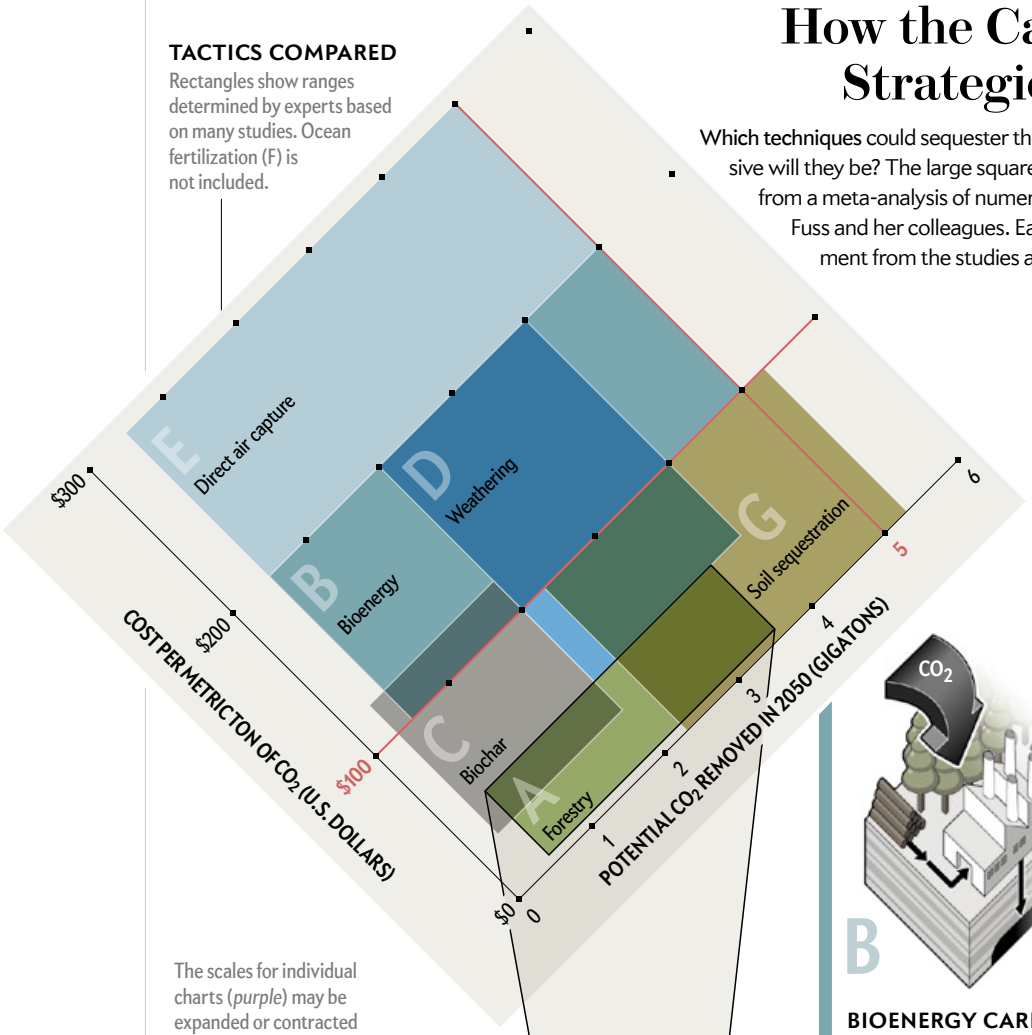
© 2019 Scientific American

How the Carbon Capture Strategies Stack Up

Which techniques could sequester the most carbon dioxide in 2050? How expensive will they be? The large square compares approaches. The numbers come from a meta-analysis of numerous studies, performed by economist Sabine Fuss and her colleagues. Each breakout chart shows a detailed assessment from the studies and the expert judgment of Fuss's team.

TACTICS COMPARED

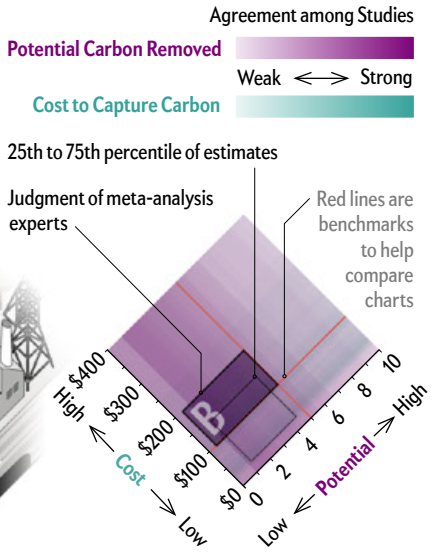
Rectangles show ranges determined by experts based on many studies. Ocean fertilization (F) is not included.



The scales for individual charts (*purple*) may be expanded or contracted for easier reading.

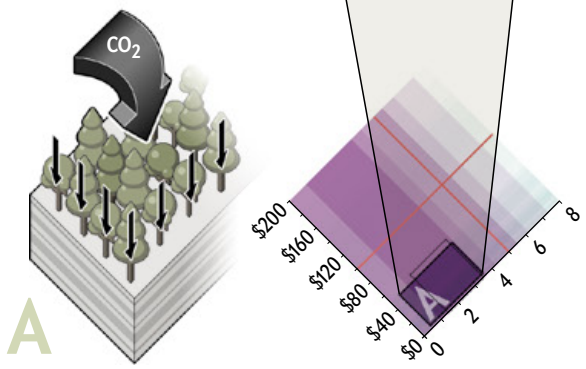
HOW TO READ THE PURPLE CHARTS

Colors depict the range of values from numerous studies. Darker shades indicate heavier overlap.



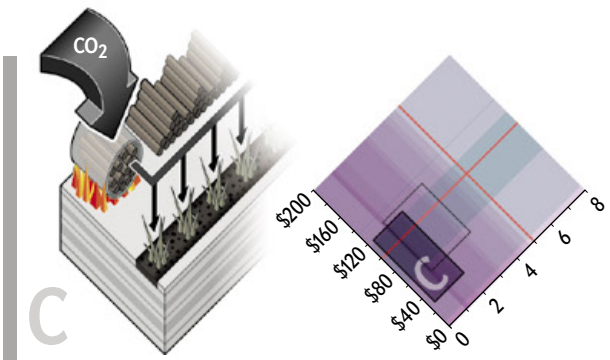
BIOENERGY CARBON CAPTURE AND STORAGE

Plants, which breathe in CO_2, are burned to produce energy or fermented into fuel. The CO_2 is removed and pumped deep underground for permanent storage. The products create revenue, but wide adoption could eat up land needed for food crops.



REFORESTATION AND AFFORESTATION

Trees are planted to replace clear-cut forests or expand existing ones. They absorb CO_2 from the air and convert it into new wood growth, including roots. Timber markets and management practices would have to be reformed.

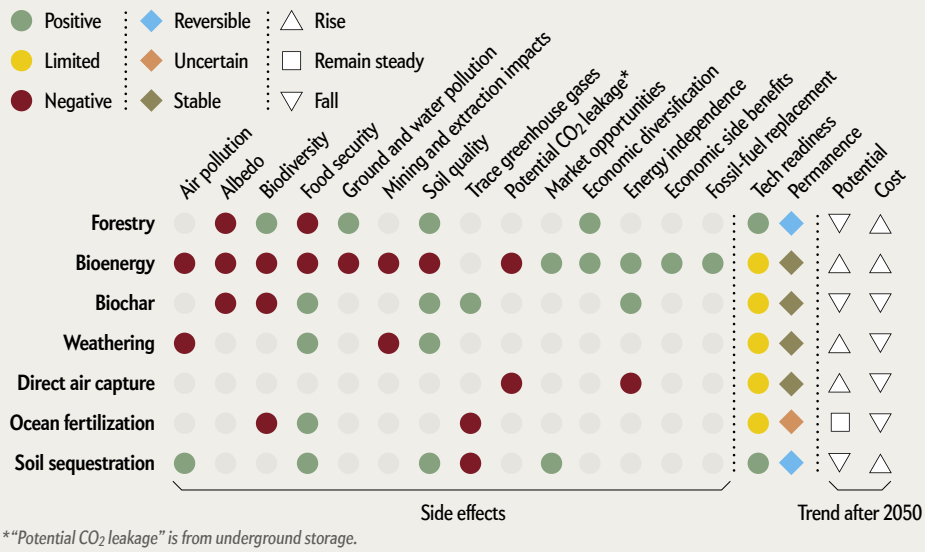


BIOCHAR

Crops, manure or organic waste is heated, without oxygen, creating biofuel and biochar—a charcoal-like residue rich in carbon. It is spread onto agricultural fields to improve soil, which can also bind additional carbon. Large-scale production with minimal energy inputs could be a challenge.

Graphics by Pitch Interactive, Illustrations by Ben Gilliland

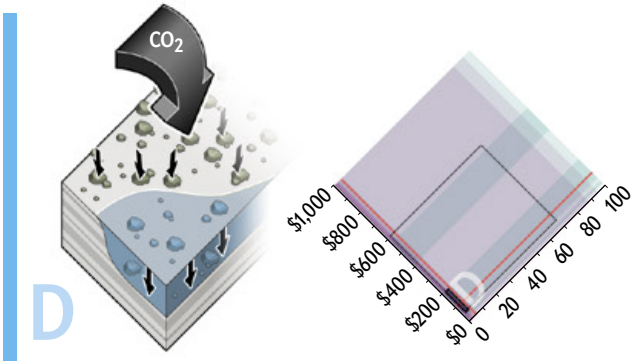
© 2019 Scientific American



*"Potential CO_2 leakage" is from underground storage.

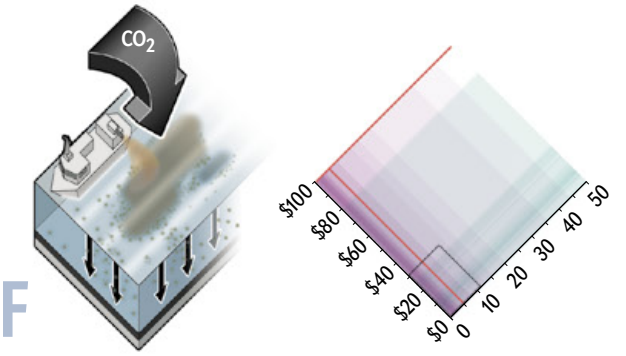
PROS AND CONS

Techniques that capture carbon in 2050 (*rows*) may come with positive or negative side effects (*dots*). The "tech readiness" column indicates how set for deployment an approach may be at that time. "Permanence" is whether stored CO_2 might escape back into the atmosphere: "reversible" indicates high risk; "stable" indicates low risk (*diamonds*). Whether the potential for sequestration or costs will rise or fall beyond 2050 is shown at the far right (*triangles*), based on factors such as land limitations. Effects are judged on how they are predominately assessed in published papers (if few papers exist, entries may be blank).



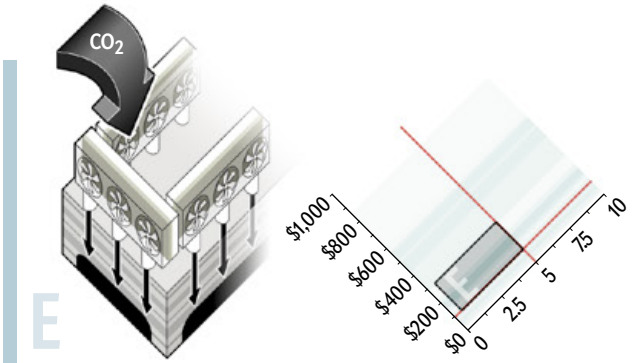
ENHANCED WEATHERING

Rock is pulverized into dust. When spread on fields, it draws CO_2 from the air and fertilizes soil. When sprinkled on the ocean, it reacts with seawater, converting CO_2 into carbonates that fall to the seafloor. Grinding and transporting rock economically are key.



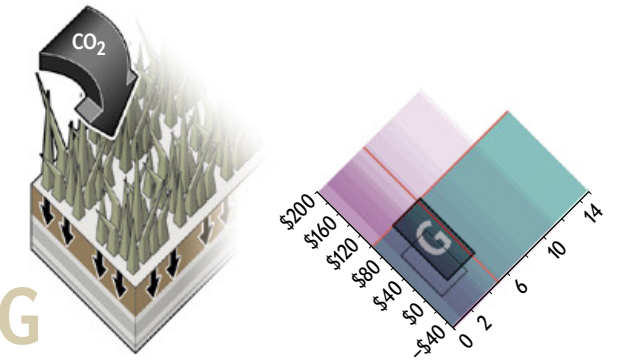
OCEAN FERTILIZATION

Iron filings are sprinkled at sea, helping plankton grow. They breathe in CO_2 and convert it into sugar or cellular material. When they die, they sink to the seafloor. Gains would be short-lived, and altering the ecosystem would be risky, so experts dropped it from the composite chart (*opposite page*).



DIRECT AIR CAPTURE

Machines pull in ambient air and chemically separate out CO_2, pumping it underground for permanent storage. Projected costs are high but could decline with technology development. (Data for "potential" were not available.)



SOIL CARBON SEQUESTRATION

Grasses or crops breathe in CO_2 and convert it into root material, fixing carbon in the soil. Soil-management techniques could enhance sequestration and plant yield. Long-term potential may be limited because soils can hold only so much carbon.

SOURCE: "NEGATIVE EMISSIONS—PART 2: COSTS, POTENTIALS AND SIDE EFFECTS," BY SABINE FUSS ET AL., IN *ENVIRONMENTAL RESEARCH LETTERS*, VOL. 13, NO. 6, ARTICLE NO. 063002; JUNE 2018

© 2019 Scientific American

would barely make a dent in the target of 20 billion tons a year.

Constructing such plants now might be essential to developing the know-how for building them on a much larger, more efficient scale later in the century. "But if you went out and built 20 million tons of direct air capture today, that would be the wrong thing to do with your money," says Roger Aines, chief scientist of Lawrence Livermore National Laboratory's energy program. "It would take a lot of solar and wind power, and if you had that much solar and wind power, the best thing to do would be to put it on the grid and turn off a coal plant." Preventing new emissions is still the overwhelming priority.

PORTFOLIO OF FIXES

THE FUSS STUDY DOES NOT simply add up the potential of the seven carbon capture methods, because some of them compete for the same resources. For example, too much reforestation would take away land needed to grow the fuel for bioenergy power plants, and too much bioenergy might compete with direct air capture for underground carbon sequestration. Climate scientists say we need to optimize a portfolio of methods.

One immediate way to start building this portfolio, says Pete Smith, a professor of soils and global change at the University of Aberdeen in Scotland, is by scaling up "stuff we already know how to do. We know how to plant trees. We know how to restore peatlands, basically by raising the water table," so the peat captures carbon dioxide instead of emitting it. "We know how to improve soil carbon content.... Incentivizing those sorts of things is [relatively] easy and could be done immediately. It would get us some of the way there."

Consider reforestation. Tragically, the world's tropical forests have become a source, rather than a sink, for carbon dioxide emissions, as trees are cut down and burned or as ravaged forests degrade. Getting forests back into negative emissions territory would first require major reforms in an international timber market that is heavily influenced by illegal trade. Beyond that, the obvious place for reforestation is land that had been clear-cut for farming or grazing but was abandoned as unproductive. Restoring five million square kilometers of such land could sequester 3.7 billion tons of carbon dioxide a year if adequate funding were available, according to a 2015 study in *Nature Climate Change* led by Richard Houghton of the Woods Hole Research Center.

Turning all livestock grazing lands that used to be forested back into forests again could create as much as 10 billion tons of negative emissions a year, according to Bronson W. Griscom, director of forest carbon science at the Nature Conservancy and lead author of a study on "natural climate solutions" in *PNAS*. That is a sizable portion of the annual carbon dioxide recovery needed annually. But the move would require a global shift away from meat eating, the opposite of the current trend.

Fuss and her co-authors foresee more modest potential. Trees live and die, meaning they will store carbon now but give it up again later this century or the next. The amount of carbon dioxide sequestered will also likely decline as forests mature, growing more slowly. Wildfires, deforestation and climate change raise the risks. Even so, forest expansion could provide a critical stopgap while direct air capture or other technologies are scaling up. Fuss puts the potential somewhere between 500 million and 3.6 billion tons of carbon dioxide removal a year by midcentury. That could take 25 billion to 180 billion tons off our target of one trillion tons this century, at a cost of $5 to $50 per ton.

BIOCHAR fertilizer was made by heating chicken waste and wood chips that would otherwise emit carbon dioxide as they decay.

JEFF HUTCHENS *Getty Images*

Better management could raise the gains. Griscom notes, for example, that tree plantation managers in the southeastern U.S. knowingly harvest loblolly pine trees several years before their optimal yield. Allowing them to sell carbon credits to cover the extra years of growth could delay harvesting to the optimal age, adding on more timber and more carbon storage.

Likewise, growing nitrogen-fixing plants in pastures and moving to a smarter system of pasture rotation could make grazing more productive while improving carbon storage in soils. Fuss conservatively estimates that improvements in soil sequestration can yield up to 5.3 billion tons a year—265 billion tons this century—at $0 to $100 per ton.

That would be in addition to biochar. In this form of carbon removal, a specialized furnace applies heat in the absence of oxygen to biomass, turning it into a form of charcoal and generating useful by-products such as bio-oil or synthetic gas. When the charcoal is applied to farm fields, it binds carbon in the soil and can improve crop yields. But no one has yet attempted to deploy biochar on a large scale. Fuss and her co-authors consider it a plausible source of 300 million to two billion tons of annual carbon dioxide removal, at $90 to $120 per ton. That is 15 billion to 100 billion tons in this century.

Another land-based approach is called bioenergy with carbon capture and storage, or BECCS. Early plans from many countries for meeting their Paris commitments depend on it, yet it is deeply controversial. A power plant burns wood, agricultural wastes or other biomass such as switchgrass. These sources take carbon dioxide out of the atmosphere as they grow or accumulate. Combustion releases it again, and the power plant recaptures it from the smokestack, sending it down into deep geologic formations for permanent storage. But reforestation for biofuel production at the scale suggested by some proponents could eat up much of the world's arable land, threatening food production and nature conservation, as well as carbon dioxide removal by other methods such as reforestation or soil sequestration. Pulling emissions from the smokestack also sharply reduces power plant efficiency, at least with current technologies. Hence, Fuss puts the sustainable yield from BECCS at just two billion tons a year, well below

© 2019 Scientific American

the forecast by other researchers, at a cost of $100 to $200 per ton. Fuss's estimate would account for 100 billion tons of negative emissions by 2100.

That leaves two other carbon capture methods now under consideration. Enhanced weathering exploits a natural process: carbon dioxide in the air converts into carbonate when exposed to certain kinds of crushed rock. The question is whether researchers can find a way to grind the right rocks into powdered form economically, to speed up the natural process. Fuss puts the potential at two billion to four billion tons a year, at $50 to $200 per ton. Her team concludes that ocean fertilization—sprinkling iron or other nutrients into the ocean to stimulate growth of algae and other plankton, which take up carbon dioxide—would be too inefficient and short-lived to justify potential ill effects on ecosystems. It is "not a viable negative emissions strategy," they write.

PROFIT RATHER THAN COST

WHERE DOES THE ACCOUNTING LEAVE US? The ranges in the Fuss study add up to as little as 150 billion tons or a bit more than one trillion tons by 2100. The latter number might sound as if it solves our problem. But we cannot just add up the numbers because of conflicts between methods. What we can do, Fuss says, is manage the portfolio to take advantage of beneficial overlaps. Enhanced weathering, for instance, could be deployed on the same land being used to grow biomass for BECCS.

What all the approaches need, scientists argue, is massive investment in research and development. "This is going to be a long, hard battle," says Lawrence Livermore's Aines. But governments have been reluctant to foot the bill for negative emissions technologies because of ideological resistance to "picking winners" and because some past investments have been notorious failures. The U.S. Department of Energy, for instance, has spent huge sums of money on carbon capture projects intended to make "clean coal" power generation a reality. Southern Company abandoned the latest attempt in 2017, switching the Kemper County clean coal plant in Mississippi to natural gas after spending $7.5 billion.

A carbon tax would bypass picking winners by imposing a cost on emissions—a cost for putting garbage into the atmosphere. That would create a marketplace motivation both to reduce emissions now and to claw back past emissions later. The U.K. imposed such a tax, currently at about $25 per ton, primarily on fossil-fuel power plants, which cut coal emissions in half just from 2015 to 2016. Most governments shy away, however, seeing a tax as too drastic for economies built on fossil fuels.

With few exceptions, corporations have also been reluctant to invest in CO_2-removal technologies because, until recently, they saw no marketplace. To them, fixing the climate is a public benefit, not something from which they can earn a profit. But that may be changing because of a surprisingly bipartisan package of tax incentives approved by the U.S. Congress in early 2018. The so-called 45Q legislation significantly increases the tax credits companies can claim over the next 12 years, not just for capturing carbon dioxide and sequestering it underground—at as much as $50 per ton in tax credits—but for using CO_2 in a variety of ways.

The most controversial use is "enhanced oil recovery." An oil company purchases carbon dioxide, transports it by pipeline and injects it into depleted oil wells, pushing out extra oil it could not extract by conventional means. A climate change solution that entails producing *more* fossil fuels may sound Orwellian, and some environmental critics have attacked 45Q as just another fossil-fuel subsidy program in disguise. But enhanced oil recovery appears to reduce current emissions because the captured carbon dioxide, typically from natural gas or ethanol refineries, gets sequestered underground. Some environmentalists, such as Kurt Waltzer of the Clean Air Task Force, argue that turning carbon capture into an energy technology, rather than an emissions technology, is the first step toward broad commercial adoption of carbon dioxide removal. It turns recaptured CO_2 into a product to be bought or sold rather than simply a cost to be endured. That could be the key to eventual negative emissions.

TIME TO START

COULD A PORTFOLIO OF CARBON capture methods, taxes and markets get us to the goal of one trillion tons by 2100? The overheated summer of 2018 may have been a turning point. The American West was on fire. People on four continents experienced severe heat waves. In Japan, thousands of heatstroke victims went to the hospital in a single week. Climate scientists shook off cautious language and warned, in *PNAS,* that further warming risks tipping the planet into a "Hothouse Earth … likely to be uncontrollable and dangerous to many." In case that message was not strong enough, senior author Hans Joachim Schellnhuber, director emeritus of the Potsdam Institute for Climate Impact Research in Germany, told reporters that the cascading effects could lead to a world capable of supporting just one billion human beings, down from 7.5 billion today.

For some political leaders even now, climate change still seems shrouded in uncertainty, despite the overwhelming evidence that it is our grim present and grimmer future. The disconcerting thing about negative emissions technologies is that so much can seem uncertain even for the scientists themselves. "Everyone is talking about how it depends on what the substrate is, what part of the world you're in, what the rainfall is like there and what the temperature is," says University of Virginia ecologist Stephanie Roe, speaking about soil carbon enhancement.

Researchers also get caught up in arguing about whether any of the carbon-removal methods, much less all of them, can scale up to billions of tons annually. "There's maybe a bit too much focus in this debate about what's the eventual scale," says Brendan Jordan of the Great Plains Institute in Minneapolis. "I fear that it paralyzes us, and we really can't afford paralysis." That is, we need to start to achieve negative emissions despite uncertainties because they are trivial compared to a world in which the climate change game of musical chairs stops, and there's no room for 6.5 billion people to sit down. SA

Richard Conniff is an award-winning science writer. His books include *The Species Seekers: Heroes, Fools, and the Mad Pursuit of Life on Earth* (W. W. Norton, 2011).

MORE TO EXPLORE

Natural Climate Solutions. Bronson W. Griscom et al. in *Proceedings of the National Academy of Sciences USA,* Vol. 114, No. 44, pages 11,645–11650; October 31, 2017.
Negative Emissions—Part 2: Costs, Potentials and Side Effects. Sabine Fuss et al. in *Environmental Research Letters,* Vol. 13, No. 6, Article No. 063002; June 2018.
Technological Carbon Removal in the United States. James Mulligan et al. World Resources Institute, September 2018.

scientificamerican.com/magazine/sa

© 2019 Scientific American

THE DENGUE DEBACLE

Is a runaway immune reaction making a dengue vaccine dangerous?

By Seema Yasmin and Madhusree Mukerjee

© 2019 Scientific American

AEDES AEGYPTI **mosquito spreads several dangerous viruses, including four that cause dengue disease. Whereas a first bout of dengue is usually mild, the second one can be lethal—a peculiarity that may be creating problems for the first-ever licensed dengue vaccine.**

© 2019 Scientific American

IN DECEMBER 2015 then president Benigno Aquino III of the Philippines and others negotiated a deal with pharmaceutical company Sanofi to purchase three million doses of Dengvaxia, the first vaccine ever licensed for dengue. The plan was to give a million schoolchildren, nine years of age, three doses of the vaccine each, sparing them from the worst outcomes of dengue: shock, organ failure and death.

1

The virus comes in four varieties. All are spread by female *Aedes* mosquitoes, primarily *Aedes aegypti,* with a penchant for sucking blood during the day, when individuals are unprotected by bed nets. In the past five decades these viruses, which are related to those that cause West Nile fever, yellow fever and Zika, have spread in waves across the tropical and subtropical world, increasing dengue incidence 30-fold and affecting upward of 390 million people each year.

Not everyone infected with a dengue virus gets sick: three out of four who get bitten will have no symptoms. The rest may suffer one of three sets of symptoms: a fever that mimics many other viral illnesses; "dengue fever," which is accompanied by headache, pain behind the eyes, aching joints and bones, and, in rare cases, internal bleeding; and severe disease encompassing dengue hemorrhagic fever and dengue shock syndrome. In severe cases, plasma seeps out of capillaries, liquid pools around organs, massive internal bleeding ensues, and the brain, kidneys and liver begin to fail. Although swift hospitalization and careful case management can and do save lives, more than 20,000 people die of dengue every year. Many are children.

Dengue is scary enough that health practitioners in developing countries have been eagerly awaiting a vaccine for decades. Yet when internist Antonio Dans and pediatrician Leonila Dans, both clinical epidemiologists at the University of the Philippines Manila College of Medicine, read about Aquino's vaccination campaign in the *Philippine Star*, the first thing that struck them was the price tag. At three billion pesos ($57.5 million) for procurement alone, the Dengvaxia campaign would cost more than the entire national vaccination program for 2015, which covered pneumonia, tuberculosis, polio, diphtheria, tetanus, pertussis, measles, mumps and rubella. It would reach less than 1 percent of the country's approximately 105 million residents. And although dengue was reported to kill an average of 750 people annually in the Philippines, it was not even among the top 10 causes of mortality. Among infectious diseases, pneumonia and tuberculosis took a far heavier toll.

Perusing an interim report from researchers at Sanofi Pasteur—the vaccine division of Sanofi—on Dengvaxia's clinical trials, Dans and Dans found further cause for concern. Among Asian children two to five years old, those who had received the vaccine were seven times *more* likely than unvaccinated children to have

IN BRIEF

A mosquito-borne disease, dengue affects almost 400 million people worldwide every year. Whereas most of those affected barely notice a first dengue infection, a second one can kill.

A controversial old theory, called antibody-enhanced development (ADE), explains why a second dengue infection can be much deadlier than the first. New studies strongly support this theory.

The first-ever vaccine licensed for dengue appears to mimic an initial dengue infection, possibly exacerbating a second one. The role of ADE in driving this phenomenon remains contested.

PRECEDING PAGES: GETTY IMAGES

© 2019 Scientific American

2

CHILD IN MANAGUA, Nicaragua, yields a blood sample (*1*) for an extensive study of dengue disease. Another child (*2*) looks down his neighborhood street.

been hospitalized for serious dengue in the third year after vaccination. Close examination of the data revealed that although the vaccine was on average safer for older children, it was statistically impossible to rule out the possibility that for some kids, Dengvaxia made things worse.

In March 2016 Dans and Dans and other medical professionals wrote to then secretary of health Janette Garin, warning that the vaccine could be risky for some children and that the Philippines may not possess enough trained health care workers to monitor so many of them for possible adverse effects. A potentially safer vaccine was in the pipeline and probably worth waiting for, they reasoned.

The same month, however, the highly respected advisory group on vaccines at the World Health Organization—which provides guidance to countries on immunization policy—stated in a briefing paper on Dengvaxia that the hospitalizations of young vaccinated children, when observed over several years, were not statistically significant. "No other safety signals have been identified in any age group" older than five, it stated. A "theoretical possibility" existed that the vaccine could be risky for some children, and further research was necessary lest the issue "compromise public confidence" in the vaccine. It nonetheless "should be introduced as part of a routine immunization program in appropriate settings." These included regions where 70 percent or more of a population had already had dengue, where immunization of early adolescents could reduce hospitalizations by up to 30 percent over a period of 30 years. A subsequent position paper from the same group stated that the vaccine was safe for children age nine and older, for whom it was recommended.

In retrospect, it did not surprise Dans and Dans that the authorities chose to ignore their concerns. "It was either believe us or believe the WHO," says Antonio Dans. "If I were them, I'd believe the WHO. I mean, who were we? We were just teachers in a small medical school." Filipino authorities were apparently so confident about Dengvaxia's safety that they did not oblige Sanofi Pasteur to submit results from so-called pharmacovigilance trials that would usually test the safety of a new drug or vaccine in local conditions. The induction of a new pharmaceutical product into the national program typically took three to five years, says Anthony Leachon, a former president of the Philippine College of Physicians, but the dengue vaccination program began right away, in April 2016.

Days later came the first report of a postvaccination fatality, of a boy with congenital heart disease. Garin explained in a press briefing that the boy's death was unrelated to Dengvaxia. Dans and Dans persisted for months, however, speaking to the press and posting a brief video on Facebook that warned—on the basis of a decades-old, highly contested theory called antibody-dependent enhancement (ADE)—that if a child had never had dengue before, the vaccine might actually make a dengue infection deadlier than it normally would have been. Garin responded with her own warning: medical practitioners who engaged in "misinfor-

PAOLO HARRIS PAZ (*1* and *2*)

© 2019 Scientific American

How Antibodies May Aggravate Dengue

Four related viruses transmitted by *Aedes* mosquitoes cause dengue disease. A theory called antibody-dependent enhancement, postulated in the 1970s by Scott B. Halstead, seeks to explain why a second bout of dengue, with a different virus, is deadlier than the first. A first infection, with, say, dengue virus 1, prompts B cells to make antibodies that coat the viruses and deliver them to white blood cells such as macrophages, which trap and kill them. The B cells subsequently become quiescent, awakening after a second infection to make the same antibodies. But antibodies to dengue virus 1 cannot bind well with, for example, dengue virus 2. The antibodies still deliver the invaders to the macrophages but in a way that enables them to escape destruction and instead commandeer the macrophages. They use the subjugated cells' machinery to replicate themselves, flooding the body with viruses. These release a protein that damages blood vessels, causing serious dengue disease.

FIRST INFECTION
SECOND INFECTION
Aedes aegypti mosquito
Dengue virus 1
Dengue virus 2
On entering the body, dengue virus often attacks skin cells
Skin cell
B cell
Antibody
Antibodies attach to virus, covering and disarming it
Antibodies are unable to attach well, leaving virus partially bare
New copies of virus are released
Macrophage
Virus is trapped and destroyed
Virus escapes cell defenses and replicates

Illustration by Tami Tolpa

© 2019 Scientific American

mation" on Dengvaxia would be responsible for every death from dengue that could have been prevented by the vaccine.

There the matter rested until November 2017, when Sanofi Pasteur issued its own advisory: those who had never experienced a dengue infection should not get Dengvaxia. A month later the WHO issued fresh guidelines, recommending the vaccine only for those with a "documented past dengue infection." The Philippines halted the vaccination program that December even as parents and the press responded with fury, recriminations and further reports of children's deaths. More than 830,000 schoolchildren had been vaccinated. According to the Department of Health (DOH), as of September 2018, 154 of the vaccinated children had died of various illnesses. The vast majority of these fatalities were unrelated to the vaccine, but clinical observations or blood tests confirmed that 19 of them had been caused by dengue.

Sanofi Pasteur contends that the deaths in the Philippines could have arisen from a failure of the vaccine to protect a small fraction of those vaccinated. In contrast, some experts argue, as Dans and Dans did, that Dengvaxia mimics a prior encounter with dengue—which can prime a patient's body to respond in a dangerous way to a second dengue infection.

The controversy has not slowed down the rollout of Dengvaxia, which is currently licensed in more than 20 countries. In October 2018 the U.S. Food and Drug Administration announced that it would prioritize review of Sanofi Pasteur's application to approve Dengvaxia. That means it could be approved in the U.S., for use in dengue-endemic areas such as Puerto Rico, before the Philippines completes its investigation into the deaths of vaccinated children—and before Sanofi Pasteur publishes its final report from the six-year-long clinical trials.

A BAFFLING DISEASE

FOR MOST VIRUSES, such as measles, the second bout, if it occurs at all, is much milder than the first. For dengue, a second bout is far more likely to kill. Scientists and doctors have struggled for years to understand why this is so. In the 1950s and 1960s, when epidemics of severe dengue began to rise in Asia, they wondered if they were dealing with an altogether new infection. The dengue they were familiar with kept patients bedridden and fatigued, but this new manifestation sent them to the hospital or the morgue. Had the virus mutated? Or was the immune system to blame?

A young scientist fresh out of medical school was seeking an answer. Scott B. Halstead began to study mosquito-borne viruses in 1957, while working for the U.S. Army in Japan. He confronted his first major dengue outbreak four years later, when stationed at a military laboratory next door to the Bangkok Children's Hospital. Doctors thought the youngsters who were carried into the hospital had been poisoned; almost a quarter of them died. Halstead led the team that identified dengue as the cause of the outbreak. He went on to make a second, more baffling, discovery. Children who were infected with dengue for a second time—each time with a different dengue virus—and babies born to mothers who were immune to dengue were most at risk for severe dengue and death. No one could explain why.

In 1964 R. A. Hawkes, then a researcher at Australian National University in Canberra, found that cell cultures infected with Murray Valley encephalitis, West Nile, Japanese encephalitis or Getah viruses infected more cells when the virus was mixed with antibodies compared with the virus alone. Hawkes proposed that the antibodies were stabilizing the virus and increasing their ability to attach to cells. Independently, Halstead was wondering if much the same was happening with dengue.

To understand why two different dengue infections were needed to make the second one lethal, Halstead infected 118 monkeys with different combinations of the four dengue viruses and measured the amount of virus in their blood. In 1973 he published his results: some monkeys, which were infected a second time and with a different dengue virus, had much higher viral loads. Four years later he provided a possible explanation, calling it antibody-dependent enhancement.

Say your first infection is with the dengue virus called DENV-1. Antibodies against that virus can linger in your blood for decades, even your entire life. When you are infected a second time with a different dengue virus, say DENV-2, 3 or 4, the antibodies against DENV-1 could paradoxically accelerate the replication of the new virus inside infected cells, precipitating a potentially fatal dengue infection.

Since refined by Halstead and other researchers, the ADE mechanism goes as follows: A dengue virus is a string of ribonucleic acid enclosed in a protein capsule, which features an array of characteristic protuberances on its surface. During a first infection with dengue, the immune system's B cells make an antibody called immunoglobulin G, or IgG, which latches onto one or more of these irregularities. On attachment, the antibodies can deliver the virus to immune system cells such as macrophages. The word "phage" derives from the Greek word meaning "to eat": macrophages are literally "big eaters." They engulf the virus and digest it with enzymes. Thus, once it is bound to antibodies, the dengue virus is normally trapped and destroyed inside macrophages.

When an infection is over, some antibody-making B cells become dormant. In the event of a second infection with a different dengue virus, these cells wake up to churn out the exact same antibodies as before. Halstead postulated that some of these antibodies can still stick to the surface of the unfamiliar virus but often fail to block its most lethal protrusions—its guns, so to speak. The antibodies still deliver the intruder to macrophages but without having disarmed it. That enables the virus to immobilize the macrophage's own defense system and take over the cell, whose resources it then uses to churn out more copies of itself. The antibody's unwitting assistance helps the new dengue variety produce 1,000 times more copies of itself than if it were acting alone.

Halstead's reward for coming up with the ADE hypothesis was a mix of indifference or disbelief from his peers, he recalls. Today, at 89 years old, he is an adjunct professor at the Uniformed Services University of the Health Sciences in Bethesda, Md., where he continues to argue his case. Many dengue experts describe him as the Godfather of ADE. "Back then, I was thinking I've made a discovery that's very important," he says. "Except nobody wanted to believe ADE was real."

More than four decades later Eva Harris, a dengue expert at the University of California, Berkeley, found strong evidence that ADE was not only real but that it contributed to severe dengue disease in children. Harris had not set out to prove or disprove ADE: she was initially skeptical of the phenomenon and not all that keen on engaging in the decades-long debate. Instead her team, including statistical modeler Leah Katzelnick, was studying the ways in which dengue sickens children. That goal then led the researchers to help establish a lab in Nicaragua and to begin one

© 2019 Scientific American

of the more challenging types of scientific projects: a long-term pediatric cohort study. Harris and her associates in Managua, Nicaragua's capital city, had the not so easy task of following thousands of children.

For more than 15 years the scientists working on the Nicaraguan Pediatric Dengue Cohort Study cared for the children if they got sick and went to their homes to collect data and blood samples. Out of 6,684 subjects, the researchers found 618 who had been sick with dengue and nearly four dozen who developed severe disease. Scouring more than 41,000 blood samples, taken over more than a dozen years, they made a striking discovery. Children with a specific concentration of antibodies—not low enough to be useless, not high enough to offer protection, but a concentration of antibodies in a middling range—were at a nearly eight times higher risk of acquiring dengue hemorrhagic fever and dengue shock syndrome.

ADE handily explains this finding. If the antibodies are not there to begin with or are present at very low densities, they cannot enhance a subsequent dengue infection to cause serious disease. If antibodies are present at high densities (as happens shortly after an initial infection), they somehow manage to cover any new dengue virus sufficiently to disable it, enabling macrophages to kill it. If, however, the antibody concentrations are in what Harris describes as a "danger zone"—not low and not high—they may facilitate the virus's entry into the macrophages without disarming it, thereby accelerating virus production.

Harris's *Science* paper describing these results was, in the words of Jean Lim, a virologist at the Icahn School of Medicine at Mount Sinai, a "rock star study" that swayed some of the staunchest naysayers of ADE. Her unexpected findings may also have hit on the solution to the dengue vaccine mystery.

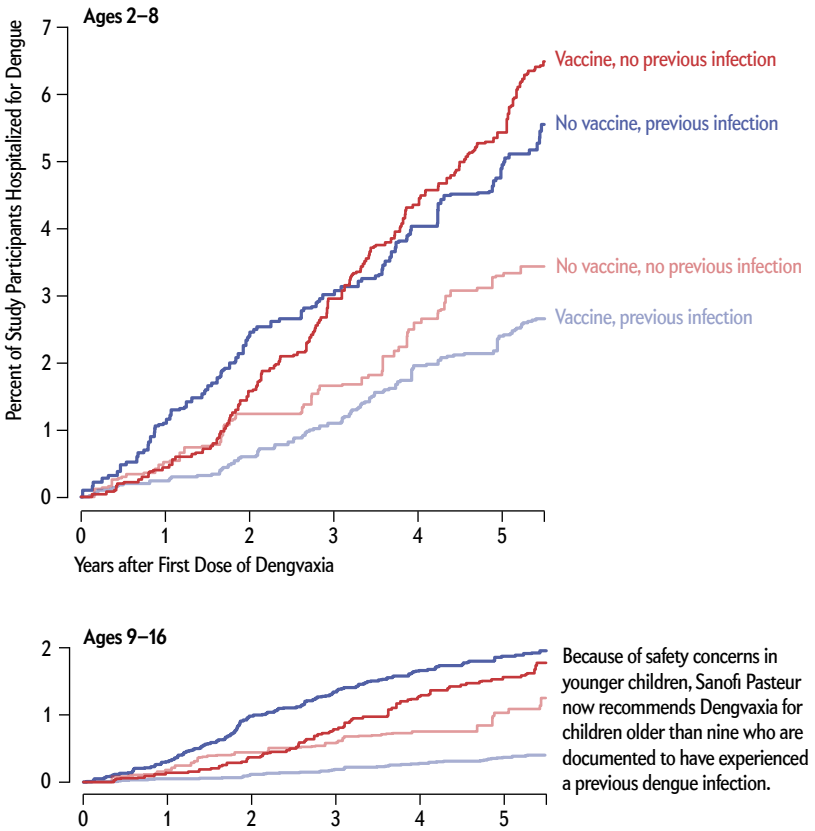
A RED FLAG

COINCIDENTALLY, DAYS AFTER Harris's paper was published in November 2017, Sanofi Pasteur made the announcement that enraged Filipino parents: do not get Dengvaxia if you have not had dengue. A month later the WHO followed suit, stating that only individuals who were proved to already have had dengue should be given the vaccine.

That was exactly what Halstead had been saying since March 2016, when he published an analysis in *Vaccine* arguing that Dengvaxia might cause harm. Perhaps in people who had never had dengue, the vaccine was acting like a first dengue infection, priming the body with just the right quantity of Trojan-horse antibodies to help a real infection turn severe. Young children were less likely to have already encountered dengue, and for them, the vaccine was more likely to act as a first infection. They were also more likely than adults to develop severe dengue after a second infection (as Halstead and others observed when a second dengue virus invaded Cuba in 1981). The problem was, there was no simple way to tell which children were dengue-negative before they received Dengvaxia—because Sanofi Pasteur had not collected those data for all of them before vaccinating them.

How Safe Is Dengvaxia?

Using a newly developed test, Sanofi Pasteur researchers evaluated which of the children in its clinical trials for Dengvaxia, the first-ever licensed vaccine for dengue, had experienced a dengue infection prior to vaccination. They found that if a child had a previous infection (*blue lines*), the vaccine was very effective in protecting him or her against hospitalization for dengue. If, however, the child had no sign of a previous dengue infection (*red lines*), he or she was far more likely than unvaccinated children of the same age group to be hospitalized with serious dengue fever, years after vaccination. The effect was far more pronounced in younger children (*top graph*), who are more likely than older children (*bottom graph*) to develop severe dengue in the first place.



Because of safety concerns in younger children, Sanofi Pasteur now recommends Dengvaxia for children older than nine who are documented to have experienced a previous dengue infection.

SOURCE: "EFFECT OF DENGUE SEROSTATUS ON DENGUE VACCINE SAFETY AND EFFICACY," BY SARANYA SRIDHAR ET AL., IN *NEW ENGLAND JOURNAL OF MEDICINE*, VOL. 379, NO. 4; JULY 26, 2018

Graphic by Amanda Montañez

© 2019 Scientific American

"I hate to say I told you so," Harris says. "But we saw this coming." At meetings and over long conference calls, she had informed Sanofi Pasteur researchers that they were not collecting the kind of data that could gauge the vaccine's potential to put lives at risk. Instead of testing all children for prior dengue infection before they received Dengvaxia, Sanofi Pasteur tested only 10 to 20 percent of them. The company argues that it was forging through unchartered territory, using the best protocols known to vaccine science. "It's routine in many vaccine trials to bleed only 10 to 20 percent of participants," says Su-Peing Ng, global medical head at Sanofi Pasteur.

After the disturbing hospitalization rate came to light, the researchers could not go back and bleed the thousands of children in the clinical trials to check their dengue status prior to vaccination. It was too late—they had already been vaccinated. Sanofi Pasteur worked with scientists at the University of Pittsburgh to develop a novel assay that could test the vaccinated children for evidence of prior dengue infection. That reassessment was the basis for the company's November 2017 warning that only those who had had dengue before should receive Dengvaxia.

The earlier recommendations had been based on the preliminary findings from the clinical trials, which showed that Dengvaxia was safer for older children. As the new tests revealed, however, age served in part as a proxy for prior dengue infection. Nine-year-olds are more likely than toddlers to have already had a dengue infection, especially in places where dengue is endemic, so giving the vaccine to them should be, on average, safe. But neither age nor endemicity is a surefire way of knowing whether a child has had dengue: the only way to know for certain is through a blood test. "Mixed in with a group of nine-year-olds will always be some kids who have never had dengue," Halstead says.

Halstead had very publicly let the WHO know about his concerns. In a December 2016 paper in the *Journal of Infectious Diseases,* he stated that a claim made by the WHO's principal advisory group on vaccines was wrong. The group had said that the risk of hospitalization for kids aged two to five peaks in the third year after vaccination and then "dissipates." Halstead argued that longer-term results from Sanofi Pasteur's clinical trials refuted this assertion. Independently analyzing the clinical trial data, Dans, Dans and others argued in a paper in the *Journal of Clinical Epidemiology* that there was "no biological basis for a threshold age of 9 years" beyond which Dengvaxia could be assumed to be safe.

The WHO stands by its decision to recommend the vaccine for older children who live in countries hardest hit by dengue, however. "The review done was extremely thorough, transparent and according to our published procedures," says Joachim Hombach, senior health adviser in the WHO's department of immunization, vaccines and biologicals. "Different options of possible recommendations were discussed, and the one published in 2016 was the consensus position of the advisory committee."

ONGOING CONTROVERSY

IN JULY 2018 Sanofi Pasteur published its reanalysis of clinical trial data using the Pittsburgh test in the *New England Journal of Medicine.* The review confirmed a higher risk of severe disease and hospitalization in "seronegative" children (those who had no evidence of prior dengue infection in their blood) who had received the vaccine, compared with those who had not. The "vaccine partially mimics primary infection and increases the risk of severe dengue during subsequent infection," the researchers wrote. Although ADE advocates had predicted this finding, the paper said that the "immunopathogenic mechanisms underlying these findings remain unknown."

Halstead contends that Sanofi Pasteur researchers are in "denial" about the evidence from their own trials. Ng counters that exactly how ADE boosts infection has yet to be demonstrated in humans. "ADE is more of a lab observation, an in vitro observation. We've not seen it clinically proven in humans," she says. "We don't know if the underlying mechanism is ADE or not." The overall impact of Dengvaxia on public health remains beneficial, Ng asserts. In children who are age nine and older and who already had dengue, Dengvaxia reduces the rate of severe disease and hospitalization by around 80 percent, according to Sanofi Pasteur. (For reasons that remain unclear, two bouts of dengue appear to confer lifelong immunity to the disease. Strictly speaking, the vaccine is useful only for those who have had one bout but not two.)

According to Sanofi Pasteur, the vaccine "partially mimics primary infection and increases the risk of severe dengue during subsequent infection."

Ng is not the only one who disputes that ADE is the main mechanism behind life-threatening dengue disease. Duane Gubler, founding chief of the dengue branch at the Centers for Disease Control and Prevention and an emeritus professor in the Emerging Infectious Diseases Program at Duke-NUS Medical School in Singapore, argues that DENV-2 and DENV-3 have historically been associated with outbreaks of severe disease. As such, the type of virus could be at least as important as ADE in determining the course of an infection. Alan Rothman, a professor of cell and molecular biology at the University of Rhode Island, says T cells, which recruit and activate macrophages and secrete inflammatory chemicals, are more directly involved in causing severe dengue than are antibodies. Halstead, on the other hand, regards T cells primarily as saviors. They kill dengue-infested macrophages, he says, at which time the viruses may release a protein that damages blood vessels. Doctors nonetheless can save a patient by maintaining his or her fluid levels, buying the T cells time to clean out the virus.

TOWARD A SAFER VACCINE

WITH DENGUE INFECTING around a million people every day and popping up in places it has never been seen before, the need for a safe vaccine is becoming ever more urgent. Armed with the new information from Sanofi Pasteur, novel dengue vaccine makers are quick to say they are doing things differently. "We've designed our trial in such a way to ask the most im-

© 2019 Scientific American

SANOFI-AQUINO-GARIN, PANAGUTIN! FILIPINO CHILDREN ARE NOT GUINEA PIGS!
SINGILIN ANG REHIMENG DUTERTE SA PAGPAPABAYA SA MGA BIKTIMA NG DENGVAXIA
GABRIELA
ILIGTAS SA PANGANIB ANG MGA BATANG BIKTIMA NG DENGVAXIA!
SAVE US
STAND FOR LIFE
DOH, NOY
AT REHIMENG DU
SA
DENGVAXI

VACCINATED CHILDREN and their parents protest the Philippines' 2016–2017 dengue immunization program.

BULLIT MARQUEZ AP Photo

portant question—how does it perform in dengue naives?" says Rajeev Venkayya, president of the Global Vaccine Business Unit at Takeda Pharmaceutical Company. Takeda is currently testing its dengue vaccine in children ages four to 16 years in Latin America and Asia. "When we started this trial in 2016, we were well aware of the concern about this issue in naives," Venkayya says. "So we made sure to have naives in our trial and collect baseline blood samples from 100 percent of participants." In January 2019 Takeda announced preliminary results from its clinical trials: the vaccine was effective. Fully assessing safety will likely take more time, however.

At least two other dengue vaccines are being developed, one by the National Institutes of Health and one by GlaxoSmithKline. They are years from being licensed—if they are found to be safe and effective. Gubler says that any vaccine will likely protect well against a couple of dengue viruses but not so well against the others. "And that being the case, there's always a risk of ADE," he continues. "So do we use those vaccines, or do we shelve them and wait another 50 years for a perfect vaccine?" Halstead is far more optimistic. "There's a really good vaccine out there," he says—the NIH vaccine, which, he wrote in a paper, "has met virtually all of the goals needed to demonstrate preclinical efficacy and safety for humans," even if it has yet to undergo extensive clinical trials.

The FDA's October 2018 announcement that it would expedite review of Dengvaxia has added fresh urgency to this debate. The burden of dengue disease in the U.S. is in territories such as Guam, the U.S. Virgin Islands, Samoa and Puerto Rico, where Gubler was based as chief of the CDC's dengue branch. He supports the use of the vaccine in places such as Puerto Rico, where, he says, the dengue surveillance system is far more robust than in the Philippines. That is, medical practitioners there should be able to keep tabs on vaccinees and ensure prompt hospitalization if they develop signs of serious disease. "I'm in favor of using it in highly endemic areas without pretesting because I think with good disease surveillance and case management, the risk of ADE is minimal," Gubler says.

Halstead disagrees: "This is a harmful product unless administered only to proven seropositive individuals." But proving previous dengue infection requires lab testing, which is not always available in many parts of the world with dengue epidemics. Controversially, the WHO advised in September 2018 that although prior screening for dengue infection was preferable, when such testing was not feasible, countries could nonetheless consider administering Dengvaxia in populations with 80 percent or higher dengue endemicity for those age nine and older. Asked to explain the ethical rationale for this recommendation, Hombach stated that the WHO had carefully weighed the pros and cons; it had also noted that such a campaign should be accompanied by "full disclosure of the risks of vaccination of persons with unknown serostatus." Effectively explaining such complex issues in ethnically diverse countries, where many people may not comprehend the languages that health officials speak or be able to read information sheets could, however, be a challenge. Sanofi Pasteur takes a more cautious view. Spokesperson Karen Batoosingh says that "the vaccine should be available for people with a prior infection to prevent against subsequent infections" and that the company is striving to develop "a new rapid dengue test to ensure broader access to the vaccine for all those who could benefit from its protective value."

© 2019 Scientific American

LOSS OF TRUST

THE REPERCUSSIONS from the vaccination program are still reverberating across the Philippines. Speaking before a senate inquiry panel, Aquino explained that dengue incidence in the country had been increasing at an alarming rate, and he had hoped that Dengvaxia could prevent the virus from invading densely populated urban areas. By this past February, however, both the senate and the house of representatives had recommended that Aquino, Garin and other senior officials be charged under an antigraft law for irregularities in the procurement and administration of the vaccine. The families of nearly three dozen dead children have brought criminal cases against Garin and other Filipino officials, accusing them of reckless imprudence amounting to homicide and torture. (Asked to comment on the circumstances in which the vaccination campaign was rolled out, Undersecretary of Health Enrique Domingo stated that he had stepped into the position in December 2017, after the uproar began, and had no personal knowledge of what had taken place.)

Amid the fear and suspicion, several outbreaks of measles have crept across the Philippines. In February 2018 the country reported that more than 8,400 have become sick and more than 130 have died. Parents were too frightened to vaccinate their kids. According to a study by the London-based Vaccine Confidence Project, in 2018 fewer than a third of Filipinos strongly agreed that vaccines are important, down from 93 percent in 2015. In that study, published in *Human Vaccines & Immunotherapeutics,* Heidi Larson, the project's director, and her co-authors argued that "biased media hype"—in particular, "false narratives aiming to vilify authorities, scientists and regulators" and "senate and congress inquiries that resembled the inquisition"—had prompted public panic and loss of trust in vaccines. Dans, Dans, Halstead and others teamed up to respond that several factors had contributed to the decline in public confidence, not least Sanofi's "exaggerated" claims of the safety of Dengvaxia: "The outrage was a result of the loss of trust rather than its cause."

Asked by *Scientific American* if he was giving ammunition to anti-vaxxers, Halstead responded that he had co-founded the Children's Vaccine Initiative in the 1990s, which later morphed into Gavi, a global public-private partnership that strives to improve vaccine access for children in poor countries. "I have very strong bona fides as a supporter of vaccines and vaccination," he says.

Even as the scientists battle it out, the parents of the vaccinated children are suffering sleepless nights, according to Antonio Dans. "The mothers are really distressed about, Was my child seronegative when he was vaccinated? Why weren't we told it could be harmful? They call us and say, My child has a cough, should we rush him to the hospital? He seems to me slightly febrile, should he go to school?" he relates. "And how do you monitor a cold and a fever in [roughly] a million kids and find out if it's dengue or not? That's a logistical nightmare, and that's what we were warning DOH about." Virtually every death in the vaccinated group was being blamed on Dengvaxia, even if it was clearly unrelated, he adds—and much of this rage and turmoil could have been avoided by accurate and timely scientific advice from trusted authorities. "So that's the sad thing here—that the WHO added to the confusion," Dans concludes.

Halstead worries that as antibody levels in the vaccinated seronegative individuals wane with time, to an intermediate level where ADE becomes more likely, they will become increasingly predisposed to developing severe dengue when they do experience an actual infection. Using Sanofi Pasteur's figures from the clinical trials—that five out of every 1,000 seronegative vaccinated children were hospitalized for dengue, of whom two had severe dengue—he calculated that more than 4,000 children could be hospitalized for vaccine-enhanced dengue disease in the Philippines. "I rub my eyes at what's happening," he says. "Why isn't Sanofi spending a lot of time thinking, 'Okay, now that we've sensitized so many people [to ADE], how are we going to protect them?'" Asked this question, Ng responded that it was unclear whether the cases of severe dengue in the vaccinated group arose from vaccine failure or ADE. All patients, regardless of whether they had dengue before or had been vaccinated or not, should guard against mosquito bites, be monitored for early signs of dengue disease, and seek prompt treatment on indications of more severe disease. Asked when the final report from the clinical trials would be published, Sanofi Pasteur responded that the outcomes had been displayed on a poster at a meeting of the American Society of Tropical Medicine and Hygiene in late 2018.

Vaccines have saved uncountable lives. Naturally occurring smallpox has been wiped off the face of the planet, and polio has almost been vanquished; tetanus and rabies no longer inspire terror. Despite these achievements, public fear of vaccines has been growing, placing millions of children at risk of avoidable disease. The increasing skepticism about vaccines is almost entirely the result of misinformation. Even so, the twists and turns of the Dengvaxia story complicate the usual narrative of valiant scientists battling public ignorance and prejudice in the quest to keep everyone safe.

The dengue saga also raises difficult questions about how pharmaceutical companies and regulators should proceed in the context of evolving scientific knowledge and imperfect vaccines. Is it ethical to endanger a minority in the interest of protecting a majority, as the WHO's September 2018 advisory on Dengvaxia implies? Who should be making these difficult decisions: global bodies of experts, national health authorities, fully informed parents and doctors, or some combination of these? And who should be held accountable when things go wrong? SA

Editors' Note: *Since this article was published in April, there have been additional deaths from measles—724 total, from all of 2018 until June 29, 2019, according to the Red Cross. In May of this year, the FDA approved Dengvaxia for use in individuals ages nine to 16. Sanofi Pasteur has still not released results from its clinical trials.*

Seema Yasmin is director of the Stanford Health Communication Initiative at Stanford University, where she also teaches science journalism and global health storytelling. She is an Emmy Award–winning reporter and author, medical doctor and frequent contributor to *Scientific American.*

Madhusree Mukerjee is *Scientific American*'s senior editor for science and society.

MORE TO EXPLORE

Dengue. Edited by Scott B. Halstead. *Tropical Medicine: Science and Practice,* Vol. 5. Imperial College Press, 2008.

Dengue Antibody-Dependent Enhancement: Knowns and Unknowns. Scott B. Halstead in *Microbiology Spectrum,* Vol. 2, No. 6, Article No. 0022-2014; December 2014.

Effect of Dengue Serostatus on Dengue Vaccine Safety and Efficacy. Saranya Sridhar et al. in *New England Journal of Medicine,* Vol. 379, No. 4, pages 327–340; July 26, 2018.

scientificamerican.com/magazine/sa

© 2019 Scientific American

© 2019 Scientific American

DIVIDE OR CONQUER

Five nations are asserting rights to vast, overlapping portions of the Arctic Ocean seafloor

By Mark Fischetti

Illustration by Peter Horvath

ON AUGUST 2, 2007, THREE RUSSIAN EXPLORERS CRAMMED inside a submersible underneath thick sea ice at the North Pole descended 4,300 meters to the dark seafloor below. They extended a robot arm from the pod and planted a titanium national flag in the sediment there. After surfacing to the supporting nuclear-powered icebreaker, expedition leader and parliament member Artur Chilingarov told an onboard reporter for the Russian news agency Itar-Tass, "If 100 or 1,000 years from now someone goes down to where we were, they will see the Russian flag." President Vladimir Putin phoned the ship, expressing his congratulations.

Canadian geophysicist David Mosher wasn't impressed when he heard the news at his Bedford Institute of Oceanography office in Nova Scotia. He glanced at a small cylinder of dried, dense mud about the size of a bratwurst lying on a plastic tray on his bookshelf. It was a short piece of a 13-meter-long sediment core extracted from the same North Pole seabed—in 1991, when Mosher was a Ph.D. student at Dalhousie University in Halifax. He had ventured there with 40 international scientists on two research icebreakers from Germany and Sweden. The scientists sent a piston corer to the seafloor, drilled down and extracted the sample from the heavy sediment.

"We didn't plant a flag," Mosher quips. "We made the hole for the Russians to plant one."

Setting the flag was a political stunt done mostly to boost the morale of Russians, who were suffer-

© 2019 Scientific American

ing through a deep recession. But the bald declaration for the North Pole made clear to the other four Arctic coastal states that the time had come to formally claim any portion of the Arctic Ocean seabed they felt they had rights to.

One of those countries was on top of it; a year earlier Norway had submitted geologic data and maps outlining three patches of seabed to the international Commission on the Limits of the Continental Shelf (CLCS), which reviews such claims and determines whether science has been applied properly. The Kingdom of Denmark, which includes Greenland, took several more years to amass a huge amount of data and submitted a pile of files in 2014, asserting it had rights to a large section of the Arctic Ocean seafloor covering 900,000 square kilometers. Russia handed in its paperwork in 2015, charting 1.3 million square kilometers—twice the size of Texas—which overlapped more than half of Denmark's outline.

This past May a Canadian team led by Mosher, who is now a geophysics professor at the University of New Hampshire, submitted 2,100 pages of text, coordinates and measurements from multibeam sonars, gravitometers and core samples to the CLCS, stating that 1.1 million square kilometers of the seabed are part of Canada. The expanse greatly overlaps the Russian and Danish claims. The U.S., the fifth state with an Arctic coastline (along Alaska), will not present its pitch until at least 2022, but its plot is expected to overlie Canada's.

For most of modern history, countries viewed the Arctic Ocean as a largely useless slab of ice. But then it started to melt, exposing opportunities. A 2008 study by the U.S. Geological Survey concluded that thick sediment in the Arctic could hold 30 percent of the world's yet to be discovered natural gas and 13 percent of its oil. Valuable iron and rare earth minerals could be waiting, too. Retreating sea ice meant shipping lanes could be opened and exploited. Seeing a bountiful future, each of the five countries became eager to secure as much territory as possible. "You never know what will happen," says Flemming Getreuer Christiansen, deputy director of the Geological Survey of Denmark and Greenland.

The CLCS could take years to work through the submissions. It moves slowly, in part because it has more than 80 cases for seafloor worldwide, from Nicaragua to Ghana to Vietnam. It is not expected to finish evaluating Denmark's or Russia's tender for several years. The Canadian review will conclude years after that. The commission does not adjudicate overlapping claims either, so once all the reviews are done countries will have to start diplomatic proceedings, putting their CLCS determinations on the table and negotiating boundary lines, another step that could take a long time.

The mapping and submission process has been civil, even cooperative, firmly grounded in geology. But the glacial pace of the process is problematic. While the scientists methodically work through the various countries' claims, Putin is expanding military bases across Russia's long Arctic shoreline. His speeches and actions have made it clear that he thinks his nation should direct the polar region. Meanwhile NATO countries are reinforcing northern militaries, wary that Russia could take over seabed the way it annexed Crimea in 2014. China is sending ships up north to signal that it, too, wants a role.

The U.S. has historically paid little attention to the region, but now it is throwing its weight around. Also in May, Secretary of State Mike Pompeo arrived at an Arctic Council meeting in Finland and declared that Russia was acting aggressively and China had to be watched closely. For the first time in 23 years, the meeting ended without the participants signing a declaration of cooperation. All the posturing could make boundary negotiations contentious, with opposing sides disregarding the science instead of compromising over it. Even worse, headstrong leaders might simply run out of patience with the CLCS's review and take what they think is theirs.

LOMONOSOV IS MINE

FOR CENTURIES nation-states saw the oceans as wild. In the 1600s they began to assert rights over the first three miles (4.8 kilometers) of seawater, based on the longest distance of a cannon shot. That practice held until the 20th century, when countries started to unilaterally claim rights out to various distances, threatening the long-standing concept of freedom of the high seas. To settle matters, in 1982 more than 160 countries agreed to the United Nations Convention on the Law of the Sea (UNCLOS). It established that a nation bordering any of the earth's oceans has an exclusive economic zone, or EEZ, reaching from its shoreline 200 nautical miles (370.4 kilometers) out to sea. It has all rights to resources in and under the water. Areas beyond that line are international waters—free to all, belonging to none.

The convention left a door open. Article 76 says a state can establish sovereign rights to exploit seabed beyond 200 nautical miles if it can present detailed geologic evidence proving that its continental shelf—the gently sloping seafloor that stretches from shore far out into the ocean before dropping into the deep sea—extends beyond the 200-nautical-mile line. Here a nation would have exclusive rights to resources on and under the seabed but not in the water column above it (fishing and navigation would remain open). Arctic countries did not focus much on the provision, until sea ice began to retreat.

Article 76 presents rules a state must follow to delineate the outer edge of a proposed extended continental shelf. It describes two formulas to draw the lines as far out as geologic evidence allows. It then describes two formulas that limit those lines, so a country cannot claim a crazy proportion of any ocean.

Both the formulas for drawing lines are based on a contour called the foot of slope. Imagine standing on the shore looking out at sea. The seafloor gradually deepens over many kilometers, then

IN BRIEF

Five countries that border the Arctic Ocean are claiming rights to large, overlapping sections of the seafloor. Three say the North Pole is theirs. Diplomats could slowly work out boundaries based on geologic evidence unless rising geopolitical tension makes the science moot.

Arctic landscapes and seascapes are changing dramatically. Rising air and water temperatures, shrinking ice and thawing permafrost are causing all kinds of living things—from algae and trees to fish and caribou—to expand their range, change migrations or, in some cases, struggle to survive.

Russia is expanding its Arctic military presence, while NATO holds large Arctic exercises, signs that aggression could mount. Yet conflict is not necessarily inevitable: countries may decide they have more to gain by cooperatively developing the changing region.

PRECEDING PAGES: GETTY IMAGES (ice)

© 2019 Scientific American

HKIRILL KUDRYAVTSEV Getty Images

STORAGE TANKS to hold liquefied natural gas are built at the Yamal LNG complex, funded in part by China and France, at the expanding Sabetta port on Russia's Arctic coast. Nations bordering the Arctic Ocean are increasingly eager to explore for gas and oil under the seafloor, and nations farther away are eyeing potentially lucrative investment opportunities.

drops down a slope to a much deeper bottom under the distant, central ocean. Along the base of the slope, scientists must determine the foot of slope—the place of maximum change in gradient—around their coastlines and islands. Generating the evidence for the foot of slope "is where all the science is," Mosher says.

Each of the Arctic Five countries, as they are known, lies along the circular rim of the pie-shaped Arctic Ocean. As they project their shelves from the perimeter toward the center, the pieces are bound to overlap: continental shelves end where plate tectonics has ended them.

Following the formulas can lead to modest overlaps, but another Article 76 provision creates a larger problem. It says a country can claim a wide band of seabed along an underwater ridge that extends from the country's continental shelf, however far the ridge goes—but it does not define what a ridge is. The language "is totally ambiguous," says Larry Mayer, director of the Center for Coastal and Ocean Mapping at the University of New Hampshire. Mayer is seen as the leading U.S. authority on Arctic seafloor and, as it happens, spent a decade as a professor at Dalhousie, where he was Mosher's Ph.D. adviser.

The ambiguity allows geologists, as well as the lawyers in their country's state department, to interpret ridge data differently. The single feature causing the greatest overlap among Denmark, Russia and Canada is the Lomonosov Ridge. It extends 1,800 kilometers from Russia's New Siberian Islands to Canada's Ellesmere Island—right next to Greenland—dividing the Arctic Ocean in half. Some of its peaks rise 3,500 meters from the deep seafloor. The ridge is a gigantic relic from millions of years ago, when the neighboring North American and Eurasian continents began pivoting away from each other, twisting and deforming the expanding Arctic Ocean floor. The ridge's common heritage means Denmark, Russia and Canada can say it naturally extends from their continental shelves, and they can outline turf along it to call their own. The most notable spot that falls within those outlines? The North Pole.

The scientists say they are just outlining where the geology takes them. But their submission teams can also apply the science to serve certain national strategies. Russia could have outlined extended continental shelf along the Lomonosov Ridge across the center of the Arctic Ocean, all the way to Canada's 200-nautical-mile EEZ, but in its submission to the CLCS it stopped just after the North Pole. It has not stated why. When I contacted two of its team experts, Eugene Petrov and Yuri Firsov, they declined to be interviewed, with Firsov e-mailing that the issues are "rather complicated." Rick Saltus, a senior research scientist at the University of Colorado Boulder, who has long been involved in the U.S. work, says Russia may not have had enough data near the Canadian end of the ridge; generating the details the CLCS looks for is expensive.

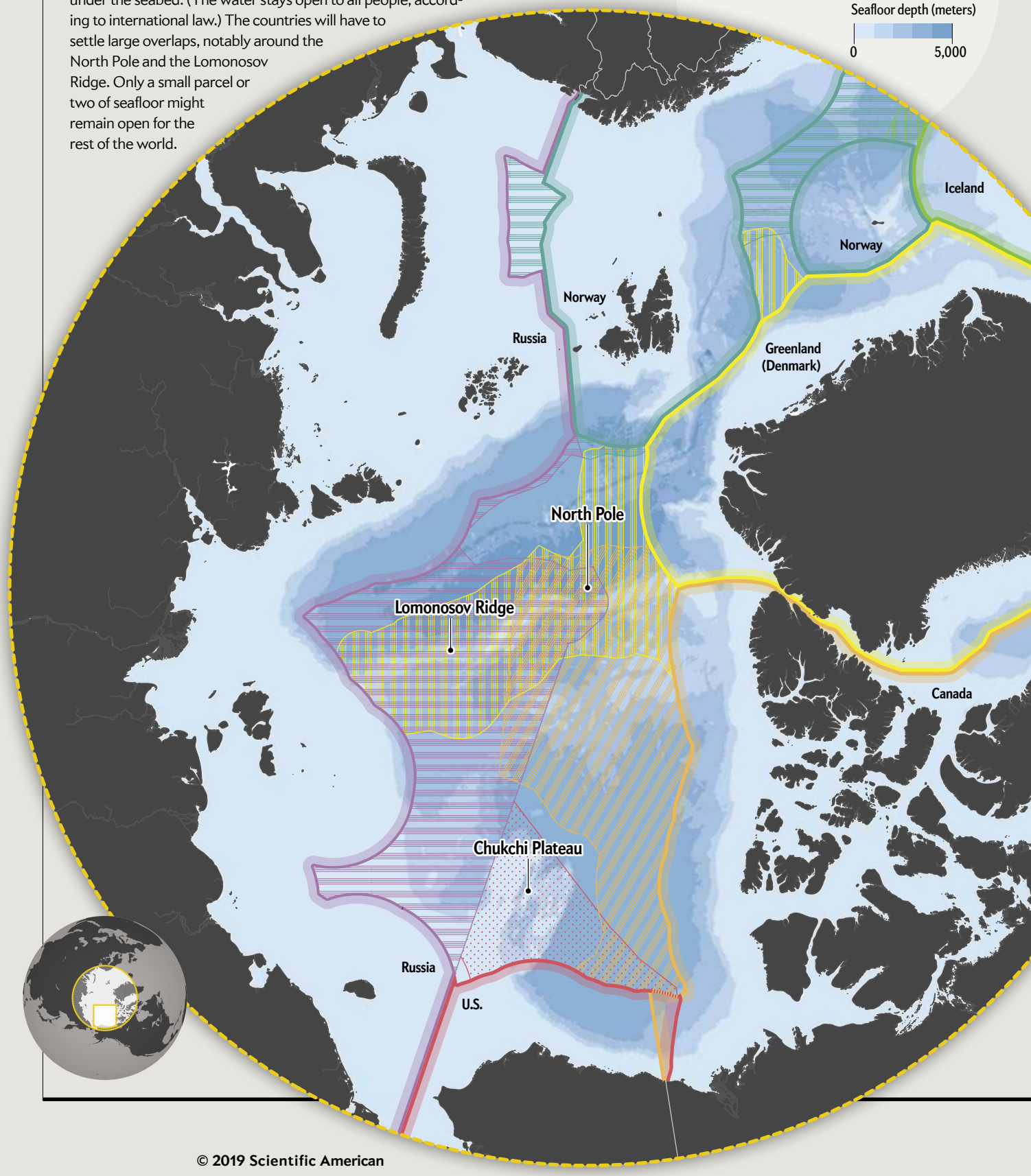
Alternatively, he says, Russia may have stopped where it did as a matter of strategy. Why complicate future boundary negotiations with Denmark and Canada? The long section of the Lomonosov Ridge that Russia did include might be more than enough to exploit.

© 2019 Scientific American

SETTING BOUNDARIES

Maps by Katie Peek, Text by Mark Fischetti

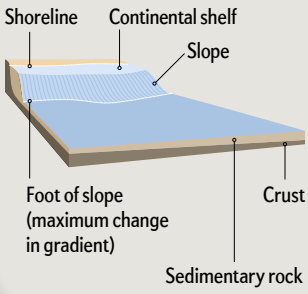
The five countries with coastlines along the Arctic Ocean are making a case to the Commission on the Limits of the Continental Shelf for "extended continental shelf"—seafloor beyond their exclusive economic zones—to gain rights to resources on and under the seabed. (The water stays open to all people, according to international law.) The countries will have to settle large overlaps, notably around the North Pole and the Lomonosov Ridge. Only a small parcel or two of seafloor might remain open for the rest of the world.



© 2019 Scientific American

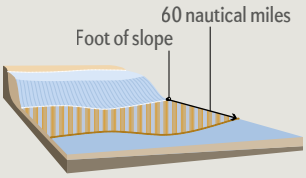
HOW COUNTRIES MAKE THEIR CASE

A country submits documents to the commission outlining the outer edge of its extended continental shelf. It uses one of two formulas (A, B) to reach as far out as they allow, starting at the foot of slope. That line is then constrained by the more lenient of two limits (C, D).



Formulas to delineate outer edge:

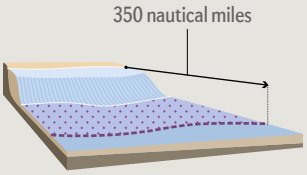
A A line 60 nautical miles from the foot of slope



B A line where the thickness of sedimentary rock is at least 1 percent of the shortest distance from the foot of slope

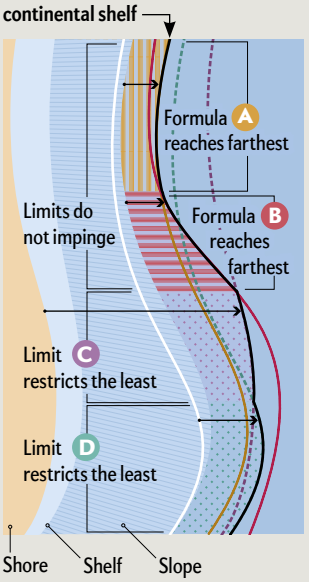
Limits to outer edge:

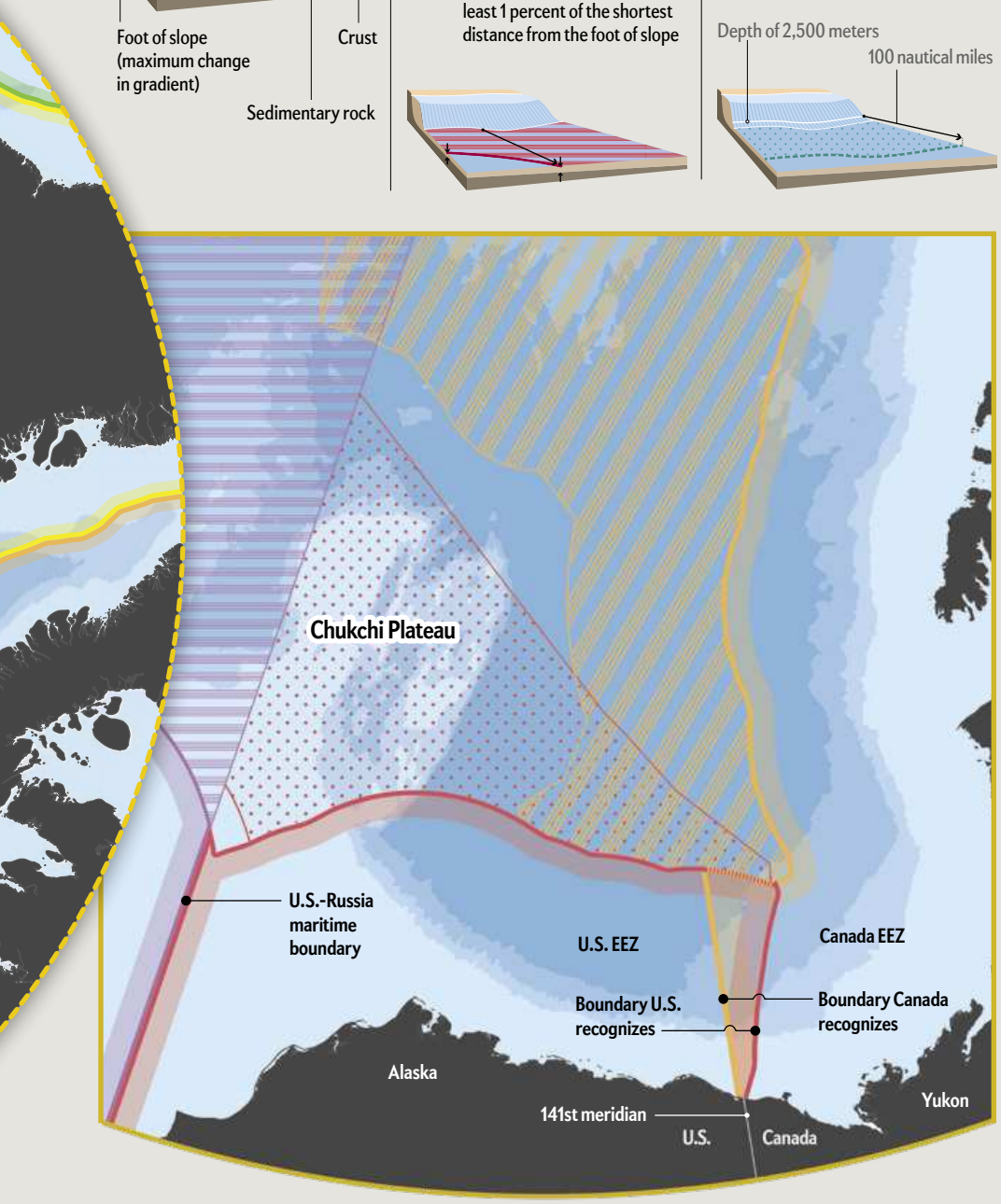
C A line 350 nautical miles from shore



D A line 100 nautical miles from a slope depth of 2,500 meters

Final outer edge of the extended continental shelf





CHUKCHI CONUNDRUM

Resolving seafloor claims involves political and scientific trade-offs.

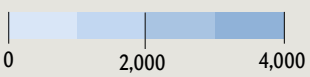
Pliable Plateau

Russia and the U.S. could say the Chukchi Plateau is a "natural prolongation" of their shelf, depending on how experts interpret the way continents have separated over millions of years. But in 1990 the former Soviet Union and the U.S. negotiated a maritime boundary between their exclusive economic zones (EEZs); Russia elongated that boundary in its submission for extended continental shelf and did not cross it. The U.S. says it will honor the boundary, too.

Economic Impasse

The U.S. and Canada disagree on their EEZ border. Canada extends the 141st meridian land boundary (*orange line*); the U.S. traces a line equidistant to the meandering coasts (*red*). The triangular seafloor in between holds an estimated 1.7 billion cubic meters of natural gas.

Seafloor depth (meters)



SOURCES: IBRU, DURHAM UNIVERSITY (*claim areas*); UNITED NATIONS CONVENTION ON THE LAW OF THE SEA (*claim areas*); MARINEREGIONS.ORG (*EEZs*); GLOBAL SELF-CONSISTENT, HIERARCHICAL, HIGH-RESOLUTION GEOGRAPHY DATABASE (*coastlines*); INTERNATIONAL BATHYMETRIC CHART OF THE ARCTIC OCEAN, VERSION 3.0 (*seafloor depths*)

© 2019 Scientific American

MELTING ICE CAPS in regions such as Svalbard (*shown*) and Greenland are exposing shorelines that can be developed, while receding sea ice exposes seafloor and shipping lanes that can be exploited throughout more of the year.

Canada has taken a similar approach, outlining the Lomonosov Ridge from its shores outward, stopping just beyond the North Pole and overlapping Russia's outline in that region. Denmark, however, claims the ridge from Greenland across the entire ocean right up to Russia's EEZ. "We are not considering whether any other states would have claims to the same area," says Denmark's lead scientist Finn Mørk, a geophysicist at the Geological Survey of Denmark and Greenland. It is up to negotiators, he says, to work out the overlaps—and who, in the end, can wave a flag from the North Pole.

POLITICS OR SCIENCE?

GIVEN THE VAGUENESS of Article 76, the three declarations for the Lomonosov Ridge might all be legitimate, scientifically. But ultimately which nation secures rights to which territory is not up to the scientists: it is up to diplomats or, potentially, militaries. And rising geopolitical tension could overtake the orderly, science-based process.

First of all, the U.S. submission to the CLCS will add to the overlaps, complicating negotiations. The extent of overlap will not be revealed until the documents are handed in, which will be 2022 at the earliest, according to Evan Bloom, director for ocean and polar affairs at the U.S. Department of State and chair of the U.S. Extended Continental Shelf Project's executive committee. Mayer says the U.S. has all the data it needs. "It's just a huge process to do the analysis," he explains.

The U.S. may be in the weakest negotiating position, however, because it has never signed the UNCLOS, unlike the rest of the Arctic Five. Plenty of U.S. officials and several presidents have recommended signing it, but a handful of treaty-wary senators have stopped the convention from ever being ratified. That may now hurt the country's own cause. "I wish the U.S. knew how much it puts itself at a disadvantage by not being a party to the convention," says Galo Carrera, a marine researcher at Dalhousie, an honorary consul of Mexico to Canada, and a former CLCS chair.

As a result, the U.S. has no need to submit a claim to the CLCS or abide by its review. But Bloom says the U.S. will do both. It has spent $89 million to obtain thorough data. It wants the rest of the world to see that it is following the same criteria as everyone else. That gives the country "very strong standing" in future negotiations, Bloom says. And there really is no other way to make a claim. The federal government could publish a document declaring "this area of seabed is ours," but the world would not recognize it. In a boundary negotiation, Saltus says, a country "would want the CLCS determination in its pocket." In effect, the U.S. recognizes the UNCLOS as customary international law—the legal practice the world follows.

U.S. rhetoric is also making the Arctic more politically complicated. In June the U.S. Department of Defense released its latest Arctic Strategy, which says that although there has been a great deal of cooperation among Arctic nations, it now anticipates an "era of strategic competition" and "a potential avenue for ... aggression."

Russia's actions could be interpreted as such. Ever since the

MICHAEL NOLAN *Getty Images*

© 2019 Scientific American

country stormed into Ukraine, "the relationship has been strained between NATO and Russia," says Rob Huebert, a political science professor at the University of Calgary and a former associate director of what is now the Center for Military, Security and Strategic Studies there. The Arctic coast offers Russia a critical strategic position for military power, notably nuclear war deterrence, because it is home to important nuclear submarine bases. "You can't separate the politics of the Arctic from the greater geopolitics" of the world, Huebert says, maintaining that Putin "sees the expansion of NATO as a core threat, and he will not allow that to happen." He says Russian jets now buzz Sweden and Finland because those countries are considering joining NATO. In March, Sweden hosted an enormous military exercise in its northernmost region with thousands of NATO troops. Because of Russia's Arctic buildup, U.S. Army General Curtis Scaparrotti told a Senate panel that same month that the U.S. military has to do more up north as well.

Russia may have another reason, beyond military strategy or oil and gas, for controlling big swaths of the Arctic seas. "It is about nationalism," says Andrew Holland, chief operating officer at the nonpartisan American Security Project. As he explains: "'We are the Russians. We are the Arctic power, and we should have rights to all of this.'"

Other geopolitical experts say the potential for Arctic confrontation is overplayed. Heather Exner-Pirot, a research fellow at the Center for Interuniversity Research on the International Relations of Canada and Quebec and an editor of the annual online *Arctic Yearbook,* which analyzes the state of Arctic politics, disagrees with Huebert, her former Ph.D. adviser. "People think there is competition in the Arctic," she says. "But what it is, really, is an oligopoly of five states that have a monopoly on the Arctic Ocean. They are thrilled with this."

The Arctic Five countries codified this situation in 2008, when they signed the Ilulissat Declaration, an agreement that says each nation will work together to safeguard marine traffic, prevent oil spills and peacefully resolve differences. It also says the countries will block any larger international attempt to govern the Arctic, as well as any other nation that might show up and try to drill for oil or gas without permission. No other countries, and no Arctic indigenous peoples, were involved.

If squabbling among members of the Arctic Five does not jeopardize orderly resolution of seafloor claims, two other wild cards could. China's economic ambitions are one of them. In 2013 President Xi Jinping unveiled the country's Belt and Road Initiative, intended to create an economic network among numerous nations by building extensive infrastructure in them all. China now heads projects in more than 60 countries worth hundreds of billions of dollars. Some world leaders worry that China's real plan is to command an enormous alliance across all of Asia. Part of the initiative is known as the Polar Silk Road, intended to develop Chinese shipping routes across the Arctic and business deals with countries along those corridors. In 2017 Xi held individual summits with the heads of Arctic nations. Not to be outdone, Putin, who has his own Eurasian vision, met one on one with leaders of Finland, Sweden, Norway and Iceland during the fifth International Arctic Forum, held in April in St. Petersburg.

The second wild card is Greenland, which has been a territory of Denmark since 1953. Although Denmark granted Greenland self-rule in 2009, the territory's 2018 political elections were a referendum supporting full independence. What has prevented secession is that the nearly 60,000 residents, who sparsely populate the largest island in the world, depend heavily on Denmark for subsidies and defense. But as ice and snow recede, China is investing in mining there. Other nations, including the U.S., are investing as well. Greenlanders are thinking they could stand on their own. Denmark has already given Greenland the seabed rights to resources within the EEZ around the island.

An independent Greenland could join NATO; the U.S. has a large air force base there. Or it could partner with China, or even Russia, to develop the thawing countryside. If Greenland becomes an independent state, Denmark could hand over the extended continental shelf claims mapped from the island's long coast. In that case, shelf negotiations might have to be recast with Greenland as the government in control—more potential delay.

END GAME

ALTHOUGH LEAD SCIENTISTS from the Arctic Five did not want to say much on the record about future boundary negotiations involving overlaps, some of them seem uneasy with the pace of the CLCS process. Nine of them—including Mosher from Canada, Mayer and Saltus from the U.S., Mørk from Denmark, and Petrov and Firsov from Russia—are working to set a common base of slope for the entire Arctic Ocean, and they are drafting a paper for a peer-reviewed journal. That would make a statement that the countries have calculated their foot of slope within the base of slope—the basis for the formulas—in the same way. Seeing such an agreement, perhaps the CLCS would speed up its reviews.

If the CLCS signs off on the Arctic Five submissions as is, only a small bit of the Arctic Ocean seabed may be left unclaimed. This space, known simply as the Area, might amount to two modest parcels far out at sea, Saltus says. The rest of the world may not be happy with that outcome. Sometimes the Arctic nations think the Arctic Ocean is their backyard, Carrera explains, but many other countries, as well as indigenous peoples, see it as a global commons. They believe they have a right to explore it for resources and to conduct research there.

Some of them think the world should formally establish the Arctic Ocean as a commons. They cite the Antarctic Treaty System as a model. In force since 1961, it sets aside all the land and ice shelves as a scientific preserve and bans military activity. It also protects more than 20 million square kilometers of the Southern Ocean around the continent. But no one lives in the Antarctic. There are no coastal states. It is more remote and more frozen. There is little insight about resources, and it offers no strategic advantage. As the Arctic warms, the once solitary home of indigenous peoples who lived off its wildness instead of trying to master it will be diced up and developed like the rest of the world to its south. Whether science or politics drives that development, it is underway. SA

Mark Fischetti is a senior editor at *Scientific American*. He covers all aspects of sustainability.

MORE TO EXPLORE

Geologic Structures of the Arctic Basin. Edited by Alexey Piskarev, Victor Poselov and Valery Kaminsky. Springer International Publishing, 2019.
United Nations Convention on the Law of the Sea: www.un.org/depts/los

scientificamerican.com/magazine/sa

© 2019 Scientific American

WHY WE TRUST LIES

The most effective misinformation starts with seeds of truth

By Cailin O'Connor and James Owen Weatherall

In the mid-1800s a caterpillar the size of a human finger began spreading across the northeastern U.S. This appearance of the tomato hornworm was followed by terrifying reports of fatal poisonings and aggressive behavior toward people. In July 1869 newspapers across the region posted warnings about the insect, reporting that a girl in Red Creek, N.Y., had been "thrown into spasms, which ended in death" after a run-in with the creature. That fall the *Syracuse Standard* printed an account from one Dr. Fuller, who had collected a particularly enormous specimen. The physician warned that the caterpillar was "as poisonous as a rattlesnake" and said he knew of three deaths linked to its venom.

Although the hornworm is a voracious eater that can strip a tomato plant in a matter of days, it is, in fact, harmless to humans. Entomologists had known the insect to be innocuous for decades when Fuller published his dramatic account, and his claims were widely mocked by experts. So why did the rumors persist even though the truth was readily available? People are social learners. We develop most of our beliefs from the testimony of trusted others such as our teachers, parents and friends. This social transmission of knowledge is at the heart of culture and science. But as the tomato hornworm story shows us, our ability has a gaping vulnerability: sometimes the ideas we spread are wrong.

IN BRIEF

Social media has facilitated the proliferation of false belief at an unprecedented scale.

By modeling the ways misinformation spreads via networks of people, researchers learn how social trust and conformity affect how communities reach consensus.

Adding propagandists to the models shows how easily belief can be manipulated, even when scientists collect ample evidence.

Over the past five years the ways in which the social transmission of knowledge can fail us have come into sharp focus. Misinformation shared on social media Web sites has fueled an epidemic of false belief, with widespread misconceptions concerning topics ranging from the prevalence of voter fraud, to whether the Sandy Hook school shooting was staged, to whether vaccines are safe. The same basic mechanisms that spread fear about the tomato hornworm have now intensified—and, in some cases, led to—a profound public mistrust of basic societal institutions. One consequence is the largest measles outbreak in a generation.

"Misinformation" may seem like a misnomer here. After all, many of today's most damaging false beliefs are initially driven by acts of propaganda and disinformation, which are deliberately deceptive and intended to cause harm. But part of what makes disinformation so effective in an age of social media is the fact that people who are exposed to it share it widely among friends and peers who trust them, with no intention of misleading anyone. Social media transforms disinformation into misinformation.

Many communication theorists and social scientists have tried to understand how false beliefs persist by modeling the spread of ideas as a contagion. Employing mathematical models involves simulating a simplified representation of human social interactions using a computer algorithm and then studying these simulations to learn something about the real world. In a contagion model, ideas are like viruses that go from mind to mind. You start with a network, which consists of nodes, representing individuals, and edges, which represent social connections. You seed an idea in one "mind" and see how it spreads under various assumptions about when transmission will occur.

Contagion models are extremely simple but have been used to explain surprising patterns of behavior, such as the epidemic of suicide that reportedly swept through Europe after publication of Goethe's *The Sorrows of Young Werther* in 1774 or when dozens of U.S. textile workers in 1962 reported suffering from nausea and numbness after being bitten by an imaginary insect. They can also explain how some false beliefs propagate on the Internet. Before the last U.S. presidential election, an image of a young Donald Trump appeared on Facebook. It included a quote, attributed to a 1998 interview in *People* magazine, saying that if Trump ever ran for president, it would be as a Republican because the party is made up of "the dumbest group of voters." Although it is unclear who "patient zero" was, we know that this meme passed rapidly from profile to profile.

The meme's veracity was quickly evaluated and debunked. The fact-checking Web site Snopes reported that the quote was fabricated as early as October 2015. But as with the tomato hornworm, these efforts to disseminate truth did not change how the rumors spread. One copy of the meme alone was shared more than half a million times. As new individuals shared it over the next several years, their false beliefs infected friends who observed the meme, and they, in turn, passed the false belief on to new areas of the network.

This is why many widely shared memes seem to be immune to fact-checking and debunking. Each person who shared the Trump meme simply trusted the friend who had shared it rather than

© 2019 Scientific American

checking for themselves. Putting the facts out there does not help if no one bothers to look them up. It might seem like the problem here is laziness or gullibility—and thus that the solution is merely more education or better critical thinking skills. But that is not entirely right. Sometimes false beliefs persist and spread even in communities where everyone works very hard to learn the truth by gathering and sharing evidence. In these cases, the problem is not unthinking trust. It goes far deeper than that.

TRUST THE EVIDENCE

THE FACEBOOK PAGE "Stop Mandatory Vaccination" has more than 140,000 followers. Its moderators regularly post material that is framed to serve as evidence for this community that vaccines are harmful or ineffective, including news stories, scientific papers and interviews with prominent vaccine skeptics. On other Facebook group pages, thousands of concerned parents ask and answer questions about vaccine safety, often sharing scientific papers and legal advice supporting antivaccination efforts. Participants in these online communities care very much about whether vaccines are harmful and actively try to learn the truth. Yet they come to dangerously wrong conclusions. How does this happen?

The contagion model is inadequate for answering this question. Instead we need a model that can capture cases where people form beliefs on the basis of evidence that they gather and share. It must also capture why these individuals are motivated to seek the truth in the first place. When it comes to health topics, there might be serious costs to acting on false beliefs. If vaccines are safe and effective (which they are) and parents do not vaccinate, they put their kids and immunosuppressed people at unnecessary risk. If vaccines are not safe, as the participants in these Facebook groups have concluded, then the risks go the other way. This means that figuring out what is true, and acting accordingly, matters deeply.

To better understand this behavior in our research, we drew on what is called the network epistemology framework. It was first

Illustration by Lisk Feng

© 2019 Scientific American

How Network Science Maps the Spread of Misinformation

We use network science to better understand how social connections influence the beliefs and behaviors of individuals in a social network—and especially how false beliefs can spread from person to person. Here we look at two kinds of network models that capture different ways in which ideas or beliefs spread. Each node in these models represents an individual. Each edge, or connection between the nodes, represents a social tie.

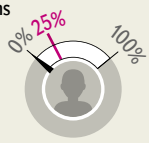
THE CONTAGION MODEL

Contagion models treat ideas or beliefs like viruses that spread between individuals in a social network. There are different ways that this "infection" can work. In some models, everyone will be infected by an infected neighbor. In others, ideas spread whenever some percentage of an individual's neighbors become infected. Here we illustrate these "complex contagions" with examples where individuals take on a new belief if at least 25 percent of their neighbors hold it. In these models, the structure of the network affects how ideas spread.

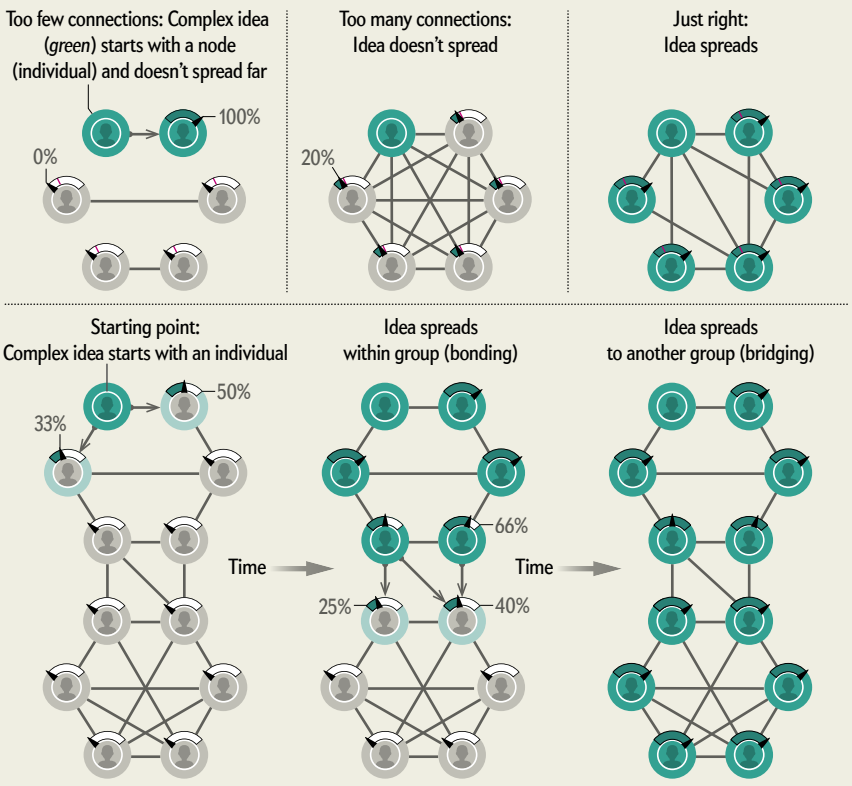
How to Read the Contagion Plots

Each circular node is a person who is influenced by ideas presented by others. Each line represents a connection between individuals.

The gauge at the top of the nodes indicates the percent of that person's connections who hold a particular belief. In the scenarios below, the threshold for an individual to take on the belief of their neighbors is 25% (at least 1 out of 4).



BONDING AND BRIDGING: In less connected groups, ideas cannot reach all members. Sometimes too many connections can also stop the spread of an idea. Some networks have tight-knit cliques, where even if an idea spreads within one clique, it can be difficult for it to spread to other cliques.



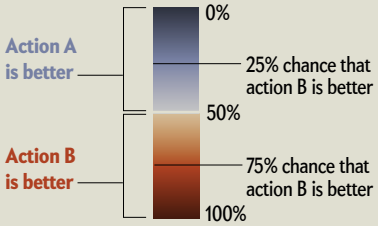
NETWORK EPISTEMOLOGY FRAMEWORK

Network epistemology models represent situations in which people form beliefs by gathering and sharing evidence. This kind of model applies to many cases in science. Beliefs do not simply spread from individual to individual. Instead each individual has some degree of certainty about an idea. This prompts the person to gather evidence in support of it, and that evidence changes beliefs. Each individual shares the evidence with network neighbors, which also influences beliefs.

How to Read the Network Epistemology Framework Plots

Each circular or square node is a person who is influenced by evidence presented by others. Each has a belief about whether **action A** (*blue*) or **action B** (*orange*) is better. The person's belief can strengthen, weaken and/or flip over time, as shown here by changing colors.

The strength of the color represents the individual's level of certainty in a particular action. For example, an assignment of 75% means that the individual thinks that there is a 75% chance that **action B** is better than **action A**. If the value is greater than 50%, then the individual performs **action B**. Then, we use Bayes's rule—which probability theory tells us is the rational way to change beliefs in light of evidence—to update the individual's credence in light of this result and then update all connections in his or her network.



Square nodes are individuals who test actions and update their beliefs accordingly (evidence seekers)

Circular nodes represent individuals who observe results from others but do not test the actions directly (observers)

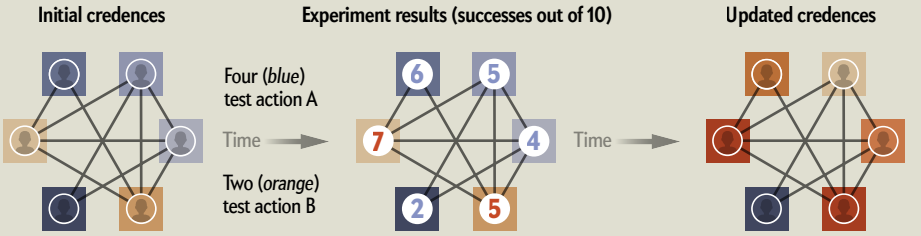
Stars represent individuals who do not hold beliefs of their own but instead focus on introducing selective results into the system (propagandists)

SOURCES: THE WISDOM AND/OR MADNESS OF CROWDS, BY NICKY CASE. *https://ncase.me/crowds*; *THE MISINFORMATION AGE: HOW FALSE BELIEFS SPREAD*, BY CAILIN O'CONNOR AND JAMES OWEN WEATHERALL. YALE UNIVERSITY PRESS, 2019

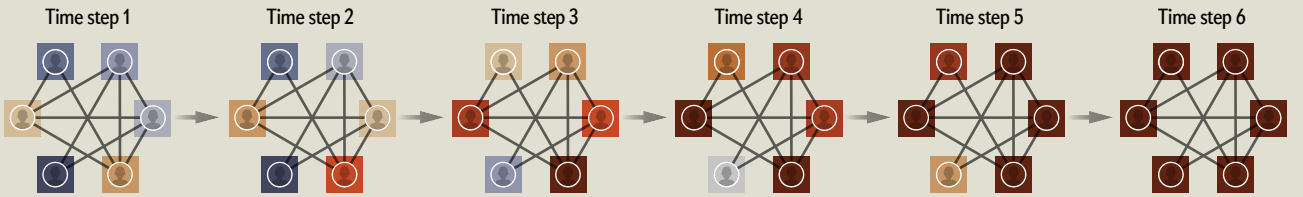
Graphic by Jen Christiansen

© 2019 Scientific American

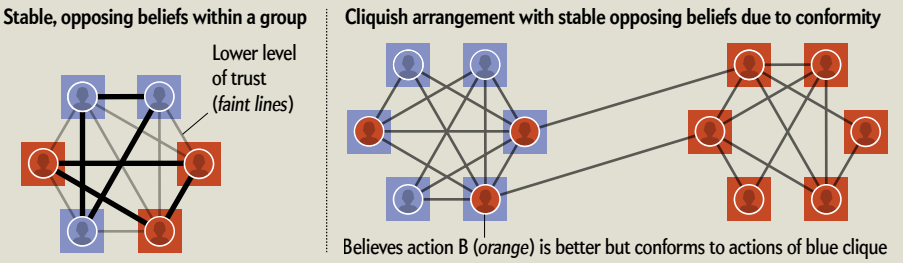
UPDATING AND EXPERIMENTING: Individuals in these models start with some random level of certainty, or credence, about whether action A or B is better. They then take the action they prefer—that is, "experiment." Their outcomes provide evidence about the success of these actions, which they share with neighbors. All individuals update their credences based on what they observe.



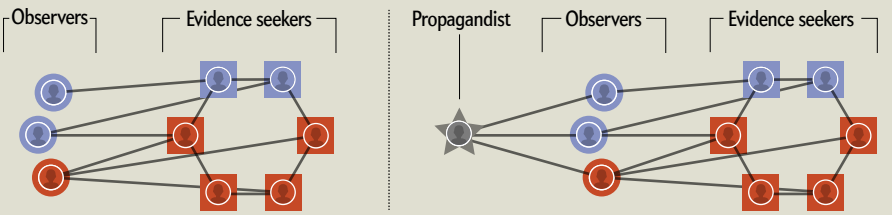
CONVERGENCE ON TRUE BELIEFS: Over time, the social connections in these models mean that groups of people come to a consensus about whether A or B is better. As they gather and share evidence, they usually learn that the better action is, indeed, better. Someone trying the worse action, for instance, will see how much better their neighbor is doing and switch. Sometimes, though, strings of misleading evidence will convince the entire group that the worse action is better.



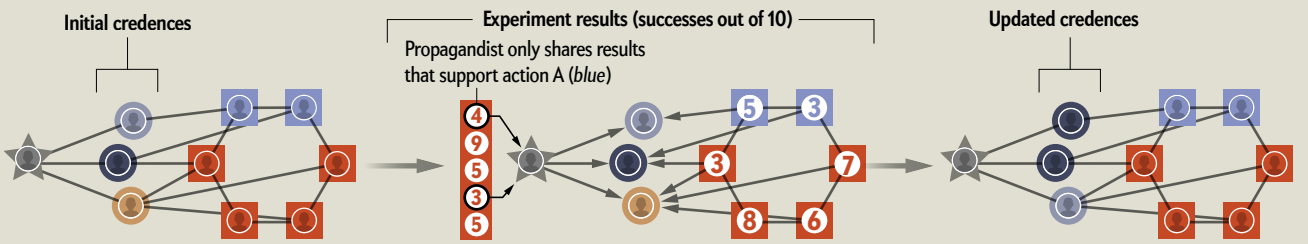
POLARIZATION: If we add social trust or conformity to these models, they may no longer reach consensus. If each individual trusts the evidence that comes from those who share their beliefs, polarized camps that only listen to those in their group form. If each individual seeks to conform their actions with group members, good ideas fail to spread between cliques.


Stable, opposing beliefs within a group
Lower level of trust (faint lines)
Cliquish arrangement with stable opposing beliefs due to conformity
Believes action B (orange) is better but conforms to actions of blue clique

EVIDENCE SEEKERS, OBSERVERS AND PROPAGANDISTS: In some cases, propagandists try to mislead a group of people about scientific results. We can use these models to represent a set of evidence seekers who gather evidence, a group of observers who update beliefs based on this evidence, and a propagandist who misleads the observers.



BELIEF UPDATING WHEN SELECTIVE RESULTS ARE IN PLAY: Industrial propagandists shape public belief by selectively sharing only those results that happen to spuriously support the worse action. This can mislead the public, even in cases when groups of evidence seekers converge to a consensus about the true belief. This strategy for public disinformation takes advantage of the inherent randomness of scientific results to mislead.



© 2019 Scientific American

developed by economists 20 years ago to study the social spread of beliefs in a community. Models of this kind have two parts: a problem and a network of individuals (or "agents"). The problem involves picking one of two choices: These could be "vaccinate" and "don't vaccinate" your children. In the model, the agents have beliefs about which choice is better. Some believe vaccination is safe and effective, and others believe it causes autism. Agent beliefs shape their behavior—those who think vaccination is safe choose to perform vaccinations. Their behavior, in turn, shapes their beliefs. When agents vaccinate and see that nothing bad happens, they become more convinced vaccination is indeed safe.

The second part of the model is a network that represents social connections. Agents can learn not only from their own experiences of vaccinating but also from the experiences of their neighbors. Thus, an individual's community is highly important in determining what beliefs they ultimately develop.

The network epistemology framework captures some essential features missing from contagion models: individuals intentionally gather data, share data and then experience consequences for bad beliefs. The findings teach us some important lessons about the social spread of knowledge. The first thing we learn is that working together is better than working alone, because an individual facing a problem like this is likely to prematurely settle on the worse theory. For instance, he or she might observe one child who turns out to have autism after vaccination and conclude that vaccines are not safe. In a community, there tends to be some diversity in what people believe. Some test one action; some test the other. This diversity means that usually enough evidence is gathered to form good beliefs.

PROTESTERS use the language of "choice" to spread misinformation about vaccine safety.

But even this group benefit does not *guarantee* that agents learn the truth. Real scientific evidence is probabilistic, of course. For example, some nonsmokers get lung cancer, and some smokers do not get lung cancer. This means that some studies of smokers will find no connection to cancer. Relatedly, although there is no actual statistical link between vaccines and autism, some vaccinated children will be autistic. Thus, some parents observe their children developing symptoms of autism after receiving vaccinations. Strings of misleading evidence of this kind can be enough to steer an entire community wrong.

In the most basic version of this model, social influence means that communities end up at consensus. They decide either that vaccinating is safe or that it is dangerous. But this does not fit what we see in the real world. In actual communities, we see polarization—entrenched disagreement about whether or not to vaccinate. We argue that the basic model is missing two crucial ingredients: social trust and conformism.

Social trust matters to belief when individuals treat some sources of evidence as more reliable than others. This is what we see when anti-vaxxers trust evidence shared by others in their community more than evidence produced by the Centers for Disease Control and Prevention or other medical research groups. This mistrust can stem from all sorts of things, including previous negative experiences with doctors or concerns that health care or governmental institutions do not care about their best interests. In some cases, this distrust may be justified, given that there is a long history of medical researchers and clinicians ignoring legitimate issues from patients, particularly women.

Yet the net result is that anti-vaxxers do not learn from the very people who are collecting the best evidence on the subject. In versions of the model where individuals do not trust evidence from those who hold very different beliefs, we find communities polarize, and those with poor beliefs fail to learn better ones.

Conformism, meanwhile, is a preference to act in the same way as others in one's community. The urge to conform is a profound part of the human psyche and one that can lead us to take actions we know to be harmful. When we add conformism to the model, what we see is the emergence of cliques of agents who hold false beliefs. The reason is that agents connected to the outside world do not pass along information that conflicts with their group's beliefs, meaning that many members of the group never learn the truth.

Conformity can help explain why vaccine skeptics tend to cluster in certain communities. Some private and charter schools in southern California have vaccination rates in the low double digits. And rates are startlingly low among Somali immigrants in Minneapolis and Orthodox Jews in Brooklyn—two communities that have recently suffered from measles outbreaks.

Interventions into vaccine skepticism need to be sensitive to both social trust and conformity. Simply sharing new evidence with skeptics will likely not help because of trust issues. And convincing trusted community members to speak out for vaccination might be difficult because of conformism. The best approach is to find individuals who share enough in common with members of the relevant communities to establish trust. A rabbi, for instance, might be an effective vaccine ambassador in Brooklyn, whereas in southern California, you might need to get Gwyneth Paltrow involved.

Social trust and conformity can help explain why polarized beliefs can emerge in social networks. But at least in some cases, including the Somali community in Minnesota and Orthodox Jewish communities in New York, they are only part of the story. Both groups were the targets of sophisticated misinformation campaigns designed by anti-vaxxers.

INFLUENCE OPERATIONS

HOW WE VOTE, what we buy and who we acclaim all depend on what we believe about the world. As a result, there are many

RICH PEDRONCELLI/*AP Photo*

© 2019 Scientific American

wealthy, powerful groups and individuals who are interested in shaping public beliefs—including those about scientific matters of fact. There is a naive idea that when industry attempts to influence scientific belief, they do it by buying off corrupt scientists. Perhaps this happens sometimes. But a careful study of historical cases shows there are much more subtle—and arguably more effective—strategies that industry, nation states and other groups utilize. The first step in protecting ourselves from this kind of manipulation is to understand how these campaigns work.

A classic example comes from the tobacco industry, which developed new techniques in the 1950s to fight the growing consensus that smoking kills. During the 1950s and 1960s the Tobacco Institute published a bimonthly newsletter called "Tobacco and Health" that reported only scientific research suggesting tobacco was not harmful or research that emphasized uncertainty regarding the health effects of tobacco.

The pamphlets employ what we have called selective sharing. This approach involves taking real, independent scientific research and curating it, by presenting only the evidence that favors a preferred position. Using variants on the models described earlier, we have argued that selective sharing can be shockingly effective at shaping what an audience of nonscientists comes to believe about scientific matters of fact. In other words, motivated actors can use seeds of truth to create an impression of uncertainty or even convince people of false claims.

Selective sharing has been a key part of the anti-vaxxer playbook. Before the recent measles outbreak in New York, an organization calling itself Parents Educating and Advocating for Children's Health (PEACH) produced and distributed a 40-page pamphlet entitled "The Vaccine Safety Handbook." The information shared—when accurate—was highly selective, focusing on a handful of scientific studies suggesting risks associated with vaccines, with minimal consideration of the many studies that find vaccines to be safe.

The PEACH handbook was especially effective because it combined selective sharing with rhetorical strategies. It built trust with Orthodox Jews by projecting membership in their community (though published pseudonymously, at least some authors *were* members) and emphasizing concerns likely to resonate with them. It cherry-picked facts about vaccines intended to repulse its particular audience; for instance, it noted that some vaccines contain gelatin derived from pigs. Wittingly or not, the pamphlet was designed in a way that exploited social trust and conformism—the very mechanisms crucial to the creation of human knowledge.

Worse, propagandists are constantly developing ever more sophisticated methods for manipulating public belief. Over the past several years we have seen purveyors of disinformation roll out new ways of creating the impression—especially through social media conduits such as Twitter bots and paid trolls and, most recently, by hacking or copying your friends' accounts—that certain false beliefs are widely held, including by your friends and others with whom you identify. Even the PEACH creators may have encountered this kind of synthetic discourse about vaccines. According to a 2018 article in the *American Journal of Public Health*, such disinformation was distributed by accounts linked to Russian influence operations seeking to amplify American discord and weaponize a public health issue. This strategy works to change minds not through rational arguments or evidence but simply by manipulating the social spread of knowledge and belief.

The sophistication of misinformation efforts (and the highly targeted disinformation campaigns that amplify them) raises a troubling problem for democracy. Returning to the measles example, children in many states can be exempted from mandatory vaccinations on the grounds of "personal belief." This became a flash point in California in 2015 following a measles outbreak traced to unvaccinated children visiting Disneyland. Then governor Jerry Brown signed a new law, SB277, removing the exemption.

Immediately vaccine skeptics filed paperwork to put a referendum on the next state ballot to overturn the law. Had they succeeded in getting 365,880 signatures (they made it to only 233,758), the question of whether parents should be able to opt out of mandatory vaccination on the grounds of personal belief would have gone to a direct vote—the results of which would have been susceptible to precisely the kinds of disinformation campaigns that have caused vaccination rates in many communities to plummet.

Luckily, the effort failed. But the fact that hundreds of thousands of Californians supported a direct vote about a question with serious bearing on public health, where the facts are clear but widely misconstrued by certain activist groups, should give serious pause. There is a reason that we care about having policies that best reflect available evidence and are responsive to reliable new information. How do we protect public well-being when so many citizens are misled about matters of fact? Just as individuals acting on misinformation are unlikely to bring about the outcomes they desire, societies that adopt policies based on false belief are unlikely to get the results they want and expect.

The way to decide a question of scientific fact—are vaccines safe and effective?—is not to ask a community of nonexperts to vote on it, especially when they are subject to misinformation campaigns. What we need is a system that not only respects the processes and institutions of sound science as the best way we have of learning the truth about the world but also respects core democratic values that would preclude a single group, such as scientists, dictating policy.

We do not have a proposal for a system of government that can perfectly balance these competing concerns. But we think the key is to better separate two essentially different issues: What are the facts, and what should we do in light of them? Democratic ideals dictate that both require public oversight, transparency and accountability. But it is only the second—how we should make decisions given the facts—that should be up for a vote. SA

Cailin O'Connor is an associate professor of logic and philosophy of science and **James Owen Weatherall** is a professor of logic and philosophy of science at the University of California, Irvine.

They are co-authors of *The Misinformation Age: How False Beliefs Spread* (Yale University Press, 2019). Both are members of the Institute for Mathematical Behavioral Sciences.

MORE TO EXPLORE

Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate. David A. Broniatowski et al. in *American Journal of Public Health*, Vol. 108, No. 10, pages 1378–1384; October 2018.

The Wisdom and/or Madness of Crowds. Nicky Case. Interactive game for network contagion: https://ncase.me/crowds

scientificamerican.com/magazine/sa

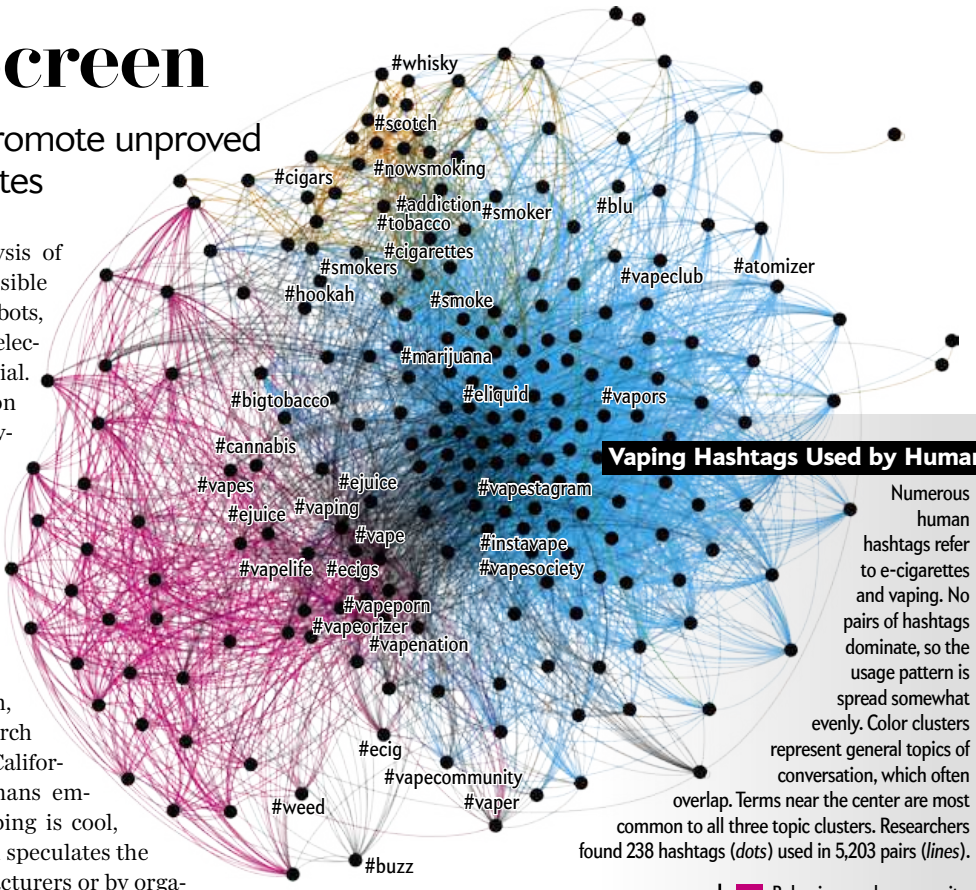
© 2019 Scientific American

Text by Mark Fischetti | Graphic by Sree Priyanka Uppu

Smoke Screen

Social media bots promote unproved benefits of e-cigarettes

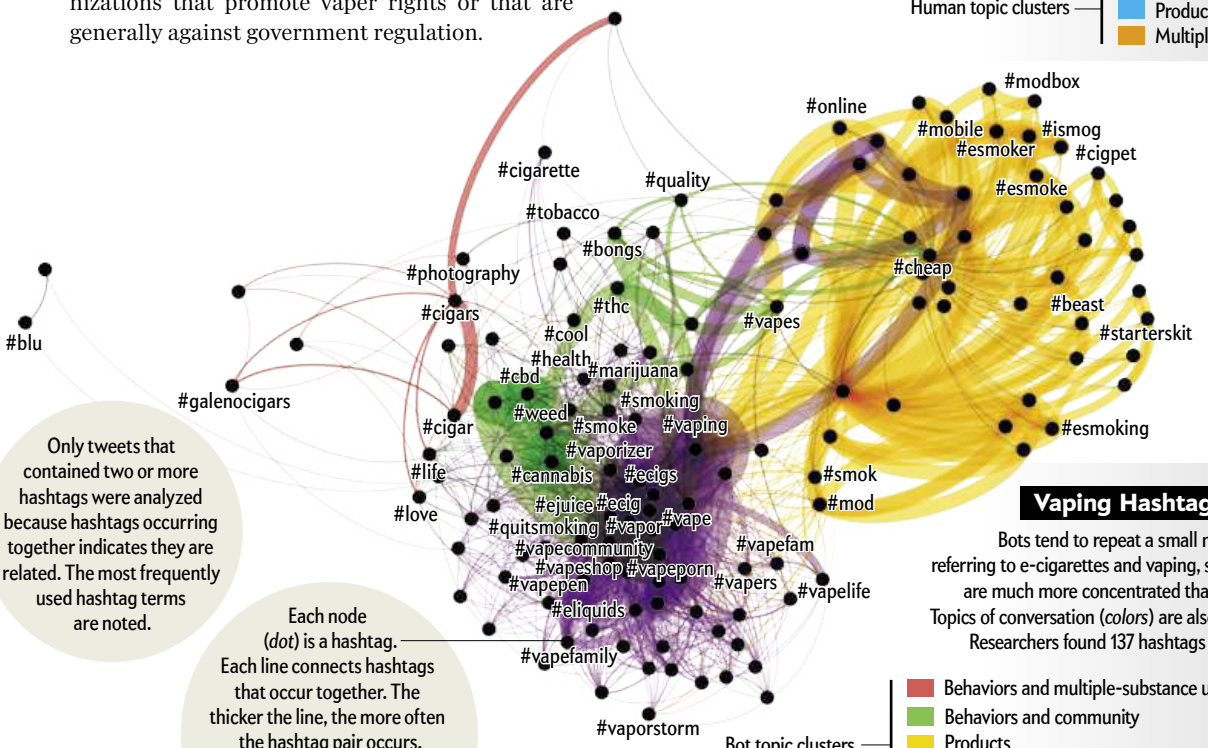
Vaping is hot. A clever analysis of Twitter posts reveals one possible reason: automated accounts, or bots, may be convincing people that electronic cigarettes are beneficial. Researchers analyzed 2.2 million tweets about vaping and discovered that hashtags used in tweets by humans differ from those in tweets by bots. Bots focus on new products and on vaping as an effective way to stop smoking tobacco even though "there is limited scientific evidence for that," says study leader Jon-Patrick Allem, an assistant professor of research at the University of Southern California. Hashtags written by humans emphasize people's lifestyles—vaping is cool, vapers are a community. Allem speculates the bots are propagated by manufacturers or by organizations that promote vaper rights or that are generally against government regulation.



Vaping Hashtags Used by Humans

Numerous human hashtags refer to e-cigarettes and vaping. No pairs of hashtags dominate, so the usage pattern is spread somewhat evenly. Color clusters represent general topics of conversation, which often overlap. Terms near the center are most common to all three topic clusters. Researchers found 238 hashtags (*dots*) used in 5,203 pairs (*lines*).

Human topic clusters
- Behaviors and community
- Products and community
- Multiple-substance use



Only tweets that contained two or more hashtags were analyzed because hashtags occurring together indicates they are related. The most frequently used hashtag terms are noted.

Each node (*dot*) is a hashtag. Each line connects hashtags that occur together. The thicker the line, the more often the hashtag pair occurs. Gray lines do not fall within the primary topic clusters.

Vaping Hashtags Used by Bots

Bots tend to repeat a small number of hashtags referring to e-cigarettes and vaping, so pairs of hashtags are much more concentrated than in human tweets. Topics of conversation (*colors*) are also more segregated. Researchers found 137 hashtags used in 1,600 pairs.

Bot topic clusters
- Behaviors and multiple-substance use
- Behaviors and community
- Products
- Health, smoking cessation and multiple-substance use

SOURCE: "E-CIGARETTE SURVEILLANCE WITH SOCIAL MEDIA DATA: SOCIAL BOTS, EMERGING TOPICS, AND TRENDS," BY JON-PATRICK ALLEM ET AL. IN *JMIR PUBLIC HEALTH AND SURVEILLANCE*, VOL. 3, NO. 4, ARTICLE E98; 2017

© 2019 Scientific American

Advertisement

CRUISE WITH SCIENTIFIC AMERICAN Travel

Cruise the Pacific Rim of South and Central America to celebrate **Scientific American's 175th Anniversary.** Savor 20+ hours of exclusive onboard classes while we're at sea. While we're in port, take advantage of archaeology, fitness, food, history, and outdoor opportunities.

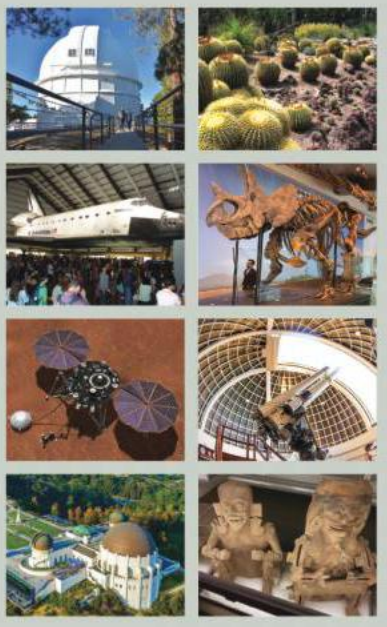
Survey the big history of the region. Get a cosmic perspective on the search for life in the universe. Enrich your knowledge of regional pre-Columbian peoples. And then head ashore and deepen your appreciation of the cultures and beauty of the area.

Cruise prices start at $2,019 per person (pp) based on double occupancy. For those attending our SEMINARS, there is a $1,575 fee. Add'l pp fees, taxes, and gratuities apply. Cruise pricing is subject to change. **For more information visit us at InsightCruises.com/events/sa39/ or email us at info@InsightCruises.com**

THE AMERICAS, MARCH 15th – 30th, 2020

Post-Cruise Land Tour March 30 – April 3, 2020

Join us in Los Angeles to visit the most exciting science sites that SoCal has to offer. We'll go behind the scenes into the prep lab at the Natural History Museum to learn how scientists process fossils. We'll also enjoy private tours at NASA's Jet Propulsion Lab, CalTech and the Mount Wilson Observatory.



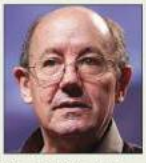
SPEAKERS & SEMINARS

The conference fee is $1,575 and includes all 90-minute seminars below.

Ken Albala, Ph.D.

ANTHROPOLOGY

- Gastronomy in the Ancient World
- The Medieval Culinary Aesthetic from Baghdad to Paris
- The Renaissance Kitchen
- Cookbooks for Mass Consumption



David Christian, Ph.D.

BIG HISTORY

- Cosmos
- A Living Planet
- Humans
- The Future: Where is it All Going?

Robert M. Hazen, Ph.D.

GEOLOGY

- How Rocks and Life Co-evolved
- Mysteries of the Evolving Mineral Realm
- Carbon and the Emergence of (Almost) Everything
- The Scientific Quest for Life's Origins

Millie Hughes-Fulford, Ph.D.

SPACE

- Living and Working in Space
- ISS and Science
- The Right Stuff — Revised 2020 Edition
- The Future

Jill Tarter, Ph.D.

HABITABLE WORLDS

- Extremophiles on Earth
- Exoplanets
- Biosignatures
- Technosignatures

For speakers' complete bios, visit http://insightcruises.com/events/sa39/#SPEAKERS.html

CST# 2065380-40

Expertise. Insights. Illumination.

Discover world-changing science. Get 12 issues of *Scientific American* in print and explore our digital archive back to 1845, including articles by more than 150 Nobel Prize winners.

sciam.com/print&archive

Download on the App Store

GET IT ON Google Play

SPECIAL ISSUE : SEX AND GENDER
GENDER MYTHS DEBUNKED
SPEAKING IN WHISTLES
WHAT'S WRONG WITH COSMIC INFLATION?
WORKOUTS AND WEIGHT LOSS
Learn the surprising evolutionary reason why exercise alon won't shed pounds—and what to do about it
PLUS
PREVENTING THREATS TO CROPS
Airborne microbes' critical role
WHAT IS A KILOGRAM?
Science seeks new ways to ensure accuracy
MEDICAL SCANS
A shortage of rare atoms imperils disease imaging
SCIENTIFIC AMERICAN
AUTOMOBILE NUME
T. REX reexamined
SCIENTIFIC AMERICAN
September 1999

Scientific American is a registered trademark of Springer Nature America, Inc. Google Play and the Google Play logo are trademarks of Google LLC. Apple, the Apple logo, iPhone, and iPad are trademarks of Apple Inc., registered in the U.S. and other countries and regions. App Store is a service mark of Apple Inc.